

**RACORO Continental Boundary Layer Cloud Investigations. Part II: Large-Eddy  
Simulations of Cumulus Clouds and Evaluation with In-Situ and Ground-Based  
Observations**

Satoshi Endo <sup>1\*</sup>, Ann M. Fridlind <sup>2</sup>, Wuyin Lin <sup>1</sup>, Andrew M. Vogelmann <sup>1</sup>,  
Tami Toto <sup>1</sup>, Andrew S. Ackerman <sup>2</sup>, Greg M. McFarquhar <sup>3</sup>, Robert C. Jackson <sup>3</sup>,  
Haflidi H. Jonsson <sup>4</sup> and Yangang Liu <sup>1</sup>

1: Brookhaven National Laboratory, Upton, NY

2: NASA Goddard Institute for Space Studies, New York, NY

3: University of Illinois, Urbana, IL

4: Naval Postgraduate School, Monterey, CA

05/05/2015

\*Corresponding author address: Satoshi Endo, Brookhaven National Laboratory,

Bldg. 815E, 75 Rutherford Drive, Upton, NY 11973, USA.

E-mail: sendo@bnl.gov

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/2014JD022525

## Abstract

A 60-hour case study of continental boundary layer cumulus clouds is examined using two large-eddy simulation (LES) models. The case is based on observations obtained during the RACORO Campaign (Routine Atmospheric Radiation Measurement [ARM] Aerial Facility [AAF] Clouds with Low Optical Water Depths [CLOWD] Optical Radiative Observations) at the ARM Climate Research Facility's Southern Great Plains site. The LES models are driven by continuous large-scale and surface forcings, and are constrained by multi-modal and temporally varying aerosol number size distribution profiles derived from aircraft observations. We compare simulated cloud macrophysical and microphysical properties with ground-based remote sensing and aircraft observations. The LES simulations capture the observed transitions of the evolving cumulus-topped boundary layers during the three daytime periods, and generally reproduce variations of droplet number concentration with liquid water content (LWC), corresponding to the gradient between the cloud centers and cloud edges at given heights. The observed LWC values fall within the range of simulated values; the observed droplet number concentrations are commonly higher than simulated, but differences remain on par with potential estimation errors in the aircraft measurements. Sensitivity studies examine the influences of bin microphysics versus bulk microphysics, aerosol advection, supersaturation treatment, and aerosol hygroscopicity. Simulated macrophysical cloud properties are found to be insensitive in this non-precipitating case, but microphysical properties are especially sensitive to bulk microphysics supersaturation treatment and aerosol hygroscopicity.

## 1. Introduction

Low-level clouds are a key component of climate and weather systems [Randall *et al.*, 2003a]. Large-eddy simulation (LES) has been widely used to investigate boundary layer cloud processes, often providing the high-resolution fields required for study that can be difficult to obtain from current observations. In particular, the boundary layer cloud working group of the Global Energy and Water Cycle Experiment (GEWEX) Cloud System Study (GCSS; currently, Global Atmospheric System Studies [GASS]) [Browning *et al.*, 1993; Randall *et al.*, 2003b] has performed many LES intercomparison studies [e.g., Bretherton *et al.*, 1999; Stevens *et al.*, 2001; Siebesma *et al.*, 2003; Stevens *et al.*, 2005; Brown *et al.*, 2002; Ackerman *et al.*, 2009; vanZanten *et al.*, 2011; Blossey *et al.*, 2013]. Their approach – idealized simulation by multiple LES models – has advanced understanding of cloud-topped boundary layers and provided reference data for single-column model (SCM) diagnostics [e.g., Zhang *et al.*, 2013].

GCSS/GASS boundary layer cloud studies to date have been limited to maritime clouds with the exception of one continental cumulus case [Brown *et al.*, 2002] based on observations from the Atmospheric Radiation Measurement (ARM) Program's Southern Great Plains (SGP) site in Oklahoma [Stokes and Schwartz, 1994; Ackerman and Stokes, 2003]. In contrast to maritime boundary layer clouds, continental boundary layer clouds typically have large temporal variations owing to the large diurnal change in surface heat flux with solar elevation. In the Brown *et al.* [2002] study, despite differences in LES dynamical schemes, the simulations agreed well with one another and with observations, and the study showed that the depth of the cloud layer is sensitive to the stability of the initial potential temperature profile. Zhu and Albrecht [2003] focused on cumulus onset in simulations using an idealized case of forced continental fair-weather cumulus clouds at the ARM SGP site. They developed a cumulus initiation diagnostic scheme based on the concept that the

strongest thermal updraft results in the deepest penetration and generates the first cumulus. They found that the development of forced cumulus is sensitive to sensible and latent heat fluxes, stability of the stratification, and moisture difference across the boundary-layer-top inversion. *Khairoutdinov and Randall* [2006] performed an idealized high-resolution simulation of shallow-to-deep cumulus transition over Amazonia, and found that updrafts lifted at the edge of the cold pool are necessary for formation of the deep cumulus circulation.

An idealized simulation approach can be used to study a simplified system by extracting a focused feature from observed cloud fields, where the uncertainty of a given simulation within a multi-model ensemble can be quantified to some degree in terms of the inter-model spread. However, more realistic configurations may be needed to reproduce observed cloud features and assess the simulations using observations. This is especially true in the case of clouds that are strongly influenced by temporally-varying forcing and aerosol properties. For example, a cloud-resolving model intercomparison study of tropical convection observed during Tropical Warm Pool International Cloud Experiment (TWP-ICE) [*Fridlind et al.*, 2012] employed time-varying forcings based on a reanalysis dataset and ARM observations [*Xie et al.*, 2004] with a multi-modal aerosol number size distribution profile as a constraint for cloud droplet activation. *Zhu et al.* [2010] alternatively used multiple nested grids to provide temporally- and spatially-varying boundary conditions for the innermost LES domain. This approach is particularly suitable for simulating inhomogeneous structures of clouds and boundary layers.

To aid understanding of continental boundary layer clouds, the Routine ARM Aerial Facility (AAF) Clouds with Low Optical Water Depths (CLOWD) Optical Radiative Observations (RACORO) campaign was held over the ARM Climate Research Facility's SGP site from April to June in 2009 [*Vogelmann et al.*, 2012]. During RACORO, in-situ aircraft observations were made around the ARM SGP site that provide aerosol

measurements for model inputs as well as in-situ and ground-based cloud observations for model evaluation. In particular, detailed cloud microphysical measurements provide data valuable for evaluation of LES microphysics. Based on observations during RACORO, three 60-hour cases of boundary layer clouds have been analyzed in the FAST-physics System Testbed and Research (FASTER) Project to study continental boundary layer cloud processes and their representation in mixed and transitional states rather than in quasi-idealized or canonical cases. This effort is detailed in Part I of this three-paper series [Vogelmann *et al.*, 2015], which develops the case studies and assesses the quality of large-scale forcing datasets.

Here in Part II we focus on the first case, in which cumulus-topped boundary layers were observed during three consecutive daytime periods. This case is selected to enable detailed examination of a single cloud type, shallow cumulus, that is not well represented in SCMs [Lin *et al.*, 2015]. Using the model input developed in Part I, we perform multiple LES simulations under continuously changing surface and large-scale forcings and aerosol conditions. The objective here is to evaluate the benchmark simulation in terms of both macrophysical and microphysical properties, through inter-model comparison and comparison with the aircraft and ground-based observations. The RACORO observations offer a unique dataset to constrain and evaluate the simulations, and comparison of the results from two LES models enables careful analysis of the impacts of different model physics treatments. In Part III of this series, these simulations are used to help diagnose parameterization biases in SCM simulations [Lin *et al.*, 2015]. We note that the continuous 60-hour period (versus three separate daytime simulations) facilitates use of the LES output as a benchmark with which to evaluate SCM simulations.

The rest of this paper is organized as follows. Section 2 describes the numerical models, model input, and experimental design. Simulated clouds are compared between models and

with the ground-based and in-situ observations in terms of cloud macrophysical properties in Section 3, and cloud microphysical properties in Section 4. The major results are summarized in Section 5.

## 2. Numerical Experiment Setup

In this section, we describe the numerical models, input forcings, and model constraints for 60-hour simulations spanning May 22, 0600 – May 24, 1800 local time (LT; UTC minus 6 h) during the 2009 RACORO field campaign. See *Vogelmann et al.* [2015] for more details on the construction of the case study and development of the model inputs.

### 2.1. Model description and configuration

Two LES models are used in this study: the Distributed Hydrodynamic Aerosol and Radiative Modeling Application (DHARMA) model [*Stevens et al.*, 2002; *Ackerman et al.*, 2000] and the Advanced Research Version of the Weather Research and Forecasting model (WRF-ARW) [*Skamarock et al.*, 2008], implemented with forcing ingestion and other functions to constitute a flexible LES in the FASTER project (WRF-FASTER) [*Blossey et al.*, 2013]. Although the two LES models used here are fewer than the number in previous intercomparisons [*e.g.*, eight models in *Brown et al.*, 2002], they do offer contrasts of the dynamic equation system and microphysics treatments as described below.

DHARMA treats fluid dynamics using an anelastic equation system [*Stevens et al.*, 2002] with a dynamic Smagorinsky sub-grid-scale scheme [*Kirkpatrick et al.*, 2006].

DHARMA has two microphysics schemes with prognostic droplet number concentration: a modified version of the *Morrison et al.* [2005] two-moment microphysics scheme, and a modified version of the Community Aerosol-Radiation-Microphysics Application (CARMA) size-resolved microphysics model [*Ackerman et al.*, 1995; *Jensen et al.*, 1998] (see Appendix

A for more information on CARMA). The CARMA spectral bin model is used primarily in a sensitivity study to examine the effect of different representations of cloud microphysics on the LES simulations. Radiative transfer is solved in DHARMA using the two-stream method as described in *Toon et al.* [1989]. In contrast to the anelastic treatment in DHARMA, WRF employs a fully compressible equation system. For this study, WRF uses a prognostic TKE turbulence scheme [*Deardorff* 1980], and a modified version of *Morrison et al.* [2005] two-moment microphysics scheme described below. Radiative transfer in WRF-FASTER is handled with the Rapid Radiative Transfer Model (RRTM) scheme [*Mlawer et al.*, 1997] adopted from the CGILS project [*Blossey et al.*, 2013]. For better representation of LES-top radiation in DHARMA and WRF, their radiative computations use a one-dimensional thermodynamic profile based on the initial sounding between the LES top and the top of atmosphere.

In the *Morrison et al.* [2005] microphysics scheme, droplets are activated from each aerosol mode as described by *Abdul-Razzak et al.* [1998], *Abdul-Razzak and Ghan* [2000], and *Ghan et al.* [2011]. Given supersaturation, both DHARMA and WRF-FASTER prognose droplet number concentration using temporally-varying multi-modal aerosol size distribution profiles and hygroscopicity parameter. In DHARMA, the activation of droplets proceeds via calculation of a prognostic saturation excess following *Morrison and Grabowski* [2008], where the supersaturation used for activation is taken as the minimum of the supersaturation over a time step (from the semi-analytic solution given in their appendix), which is found to produce close agreement with droplet concentrations computed using bin microphysics in DHARMA simulations of shallow convection. In WRF-FASTER, the activation of droplets proceeds via calculation of supersaturation from the prognostic potential temperature and water vapor mixing ratio on entry to the microphysics routine (before any adjustments from microphysical processes). Although the original Morrison et

al. scheme predicts mixing ratio and number concentration for cloud droplets, in the distributed version of WRF it only predicts mixing ratio for cloud droplets. This study uses a version that also predicts number concentration, with an extension to activate cloud droplets using temporally-varying quad-modal aerosol size distribution profiles and a hygroscopicity parameter. While WRF-FASTER treats aerosol diagnostically, DHARMA treats aerosol number concentrations prognostically, including advection and consumption, and the mean aerosol profile is nudged to time-dependent observation-derived values with a six-hour time scale.

The two models have a common configuration in certain aspects, with a domain size of  $9.6 \text{ km} \times 9.6 \text{ km}$ , horizontal grid points of  $128 \times 128$  (75-m resolution), and a vertical resolution of 40 m for the 125 levels below 5 km. The model domain is smaller than the standard ARM SGP observational domain ( $\sim 300 \text{ km} \times 300 \text{ km}$ ) but is larger than the spatial scale of shallow cumuli that are the focus of this study. A sensitivity test using a domain three times larger produces negligible change in cloud properties (not shown). Above 5 km, DHARMA uses 20 stretched grid levels up to 14.5 km, and WRF-FASTER uses a sponge layer for 13 grid levels up to 5.5 km. Although WRF employs a hydrostatic pressure vertical coordinate [ $\eta = (p_h - p_{ht})/(p_{hs} - p_{ht})$ , where  $p_h$  is hydrostatic component of pressure and, respectively,  $p_{ht}$  and  $p_{hs}$  are the values at model top and surface], the  $\eta$ -spacing is specified to provide a uniform step size in geometric height. Periodic boundary conditions are used at the lateral boundaries.

In addition to the baseline simulations using the two models with the two-moment microphysics, simulations with bin microphysics and sensitivity studies are performed using DHARMA to examine the influences of the treatment of aerosol advection, supersaturation, and activation. Table 1 summarizes the models used, experiment designs, and key features.

Briefly: DHARMA PRNA is the baseline simulation using the above configuration including



the prognostic aerosol number concentration ( $N_a$ ) treatment; DHARMA DINA uses diagnostic aerosol concentrations in place of the prognostic treatment in DHARMA PRNA; DHARMA MAXS, in addition, diagnoses grid-scale supersaturation as in WRF-FASTER. Simulations labeled as 'BIN' use the size-resolved bin microphysics scheme in place of the two-moment microphysics scheme (with substepping during droplet activation), and those labeled 'K' replace the observationally-derived aerosol hygroscopicity values ( $\sim 0.1$ ) used in other simulations with a constant value representative of pure ammonium bisulfate (0.55).

## 2.2. Model input

### 2.2.1. Aerosol size distribution parameters and hygroscopicity

One of the unique opportunities in this study is the use of the time-varying trimodal aerosol size distribution profiles and hygroscopicity parameters developed in *Vogelmann et al.* [2015]. During the 60-hour case study period, three aircraft flights yielded six vertical profiles over the ARM SGP site from spiral ascents and descents. Each flight consists of horizontal legs in a triangular pattern at different altitudes for cloud sampling, and two vertical spirals for profiling before and after the horizontal legs.

For each profile, aerosol size distributions were generated from measurements by a Scanning Mobility Particle Sizer (SMPS) and a Passive Cavity Aerosol Spectrometer Probe (PCASP). Each aerosol size distribution was fit using three lognormal-modes; to simplify implementation in simulations while sacrificing little accuracy, the mean geometric diameter and standard deviation of each mode was fixed within each vertical profile but permitted to vary from one profile to the next, whereas the number concentration was permitted to vary both vertically and temporally [see *Vogelmann et al.*, 2015]. The number concentrations from the highest and lowest levels in each profile are extended, respectively, to the surface and to the model top levels. The parameters are then interpolated over time between the six

profiles. Owing to the intermittent presence of prominent peaks across a wide size spectrum, four modes were used in practice for interpolation, wherein no more than three modes contained particles at any single time. For the periods before (after) the first (last) profile, the values from the first (last) profile are used.

Aerosol hygroscopicity was determined based on  $\kappa$ -Köhler theory [Petters and Kreidenweis, 2007] for each profile using the aerosol size distributions and the CCN measured by a Droplet Measurement Technologies (DMT) Dual-Column CCN Spectrometer. The estimated aerosol hygroscopicity is interpolated temporally in the same manner as the size distribution parameters.

Figure 1 shows time-height variation and vertical profiles of aerosol number concentration for the particles larger than  $0.03\ \mu\text{m}$  in radius, which includes the accumulation mode ( $> 0.05\ \mu\text{m}$  radius) that significantly contributes to cloud droplet activation. The aerosol number concentration is greatest in the lower atmosphere, and tends to decline with time during the case study period.

#### 2.2.2. Initial conditions

The models are initialized at 0600 LT with the vertical profiles obtained from a rawinsonde launched at 0530 LT on 22 May 2009 from the ARM SGP Central Facility. The profile has a typical early morning structure consisting of a near-surface stable layer below 400 m and a nearly neutral residual layer between 400 and 2400 m (not shown).

#### 2.2.3. Large-scale forcings

Whereas Vogelmann *et al.* [2015] examine large-scale forcings from different sources and the impact of using relaxation on mean fields, this study uses the forcing from the standard version of the ARM Continuous Variational Analysis (VARANAL) forcing product [Zhang

*et al.*, 2001; *Xie et al.*, 2004] with relaxation. The ARM forcing product provides surface and large-scale forcing based on a variational analysis that includes ARM surface observations as constraints to optimize the forcing derived from objective analysis or reanalysis data (NOAA/NCEP Rapid Updated Cycle [RUC] for the version used in this study).

The large-scale advective tendencies for potential temperature  $\theta$  and water vapor mixing ratio  $q_v$  are written as,

$$\left(\frac{\partial \bar{\theta}}{\partial t}\right)_{\text{LS}} = -\bar{\mathbf{v}} \cdot \nabla \bar{\theta} - \bar{w} \frac{\partial \bar{\theta}}{\partial z} \quad (1)$$

and,

$$\left(\frac{\partial \bar{q}_v}{\partial t}\right)_{\text{LS}} = -\bar{\mathbf{v}} \cdot \nabla \bar{q}_v - \bar{w} \frac{\partial \bar{q}_v}{\partial z} \quad (2)$$

where LS denotes ‘large scale’,  $\mathbf{v} = (u, v)$  is the horizontal wind velocity,  $w$  is the vertical wind velocity, and the overbar denotes a large-scale mean. The first and second terms on the right-hand side of the equation represent large-scale horizontal and vertical advection, respectively.

For these simulations, the effect of large-scale advection is represented by additional terms in the governing equations in the model,

$$\left(\frac{\partial \theta}{\partial t}\right)_{\text{LS}} = -\mathbf{V} \cdot \nabla \theta - W \frac{\partial \theta}{\partial z} \quad (3)$$

and,

$$\left(\frac{\partial q_v}{\partial t}\right)_{\text{LS}} = -\mathbf{V} \cdot \nabla Q_v - W \frac{\partial q_v}{\partial z} \quad (4)$$

where the capital variables ( $\theta$ ,  $Q_v$ ,  $\mathbf{V}$ , and  $W$ ) represent the large-scale fields associated with external forcing to distinguish them from model prognostic variables indicated by the lower-case variables ( $\theta$ ,  $q_v$ ,  $\mathbf{v}$  and  $w$ ). While the horizontal advection terms and  $W$  in the vertical advection term are prescribed from the VARANAL forcing dataset, we use the

vertical gradients of the state variables  $\partial\theta/\partial z$  and  $\partial q_v/\partial z$  in the simulation to compute the local effect of large-scale vertical motion.

In addition to the advective tendencies above, relaxation (nudging) to the reference profile is applied to horizontal-mean  $\theta$ ,  $q_v$ ,  $u$  and  $v$  using the following equations for scalars,

$$\left(\frac{\partial\theta}{\partial t}\right)_R = \frac{\theta - \bar{\theta}}{\tau} \text{ and } \left(\frac{\partial q_v}{\partial t}\right)_R = \frac{q_v - \bar{q}_v}{\tau} \quad (5)$$

and for momentum,

$$\left(\frac{\partial u}{\partial t}\right)_R = \frac{U - \bar{u}}{\tau} \text{ and } \left(\frac{\partial v}{\partial t}\right)_R = \frac{V - \bar{v}}{\tau} \quad (6)$$

where the subscript R denotes relaxation, and  $\tau$  is the relaxation timescale per term.

For this study, the reference values (capitals) are from the ARM forcing dataset. The inverse relaxation timescale changes with height as follows:

$$\begin{aligned} \frac{1}{\tau} &= 0 \quad (z \leq 400 \text{ m}) \\ &= \left(\frac{1}{\tau_{full}}\right) \frac{z - 400}{200} \quad (400 \text{ m} < z \leq 600 \text{ m}) \\ &= \left(\frac{1}{\tau_{full}}\right) \quad (600 \text{ m} < z) \end{aligned} \quad (7)$$

where  $\tau_{full}$  is the relaxation time scale at full strength (*i.e.*, without height variation), which is 12 h for  $\theta$  and  $q_v$ , and 3 h for  $u$  and  $v$ . The values are selected so that they are short enough to inhibit significant drift of the simulated fields, but long enough to avoid having undue influence on the evolution of the cumulus convection (see discussion in *Vogelmann et al.*, 2015). Applying relaxation to the horizontal mean variables preserves the structure in the deviations from the mean.

#### 2.2.4. Surface forcing

Time-varying surface sensible and latent heat fluxes are prescribed from the ARM forcing dataset. Surface momentum flux is computed in the surface layer scheme using the Monin-Obukhov similarity with an aerodynamic roughness length of 0.04 m, which is slightly larger than in the GCSS ARM SGP case (0.035) and is a typical value in northern central Oklahoma during May [David Cook, personal communication]. Surface albedo and emissivity are set to 0.2 and 1.0, respectively. Time variation of surface skin temperature is also prescribed from the forcing dataset for radiation.

### 3. Cloud Macrophysical Properties

During the case study period, the cumulus-topped boundary layer was observed to develop three times. Figure 2 shows the temporal variation of the cloud occurrence profile from the ARM Active Remote Sensing of Clouds (ARSCL) product [*Clothiaux et al.*, 2000], the lifting condensation level (LCL), and the surface sensible and latent heat fluxes. The ARSCL product represents cloud frequency of occurrence every 5 minutes at the ARM SGP Central Facility, based on continuous measurements from vertically-pointing cloud radar, micropulse lidar, and ceilometer instruments. The LCL is the averaged value from five surface stations closest to the flight path, which include the ARM Central Facility and Oklahoma Mesonet stations. The surface heat fluxes, taken from the ARM VARANAL forcing dataset, represent area-weighted averages of many sites within the ARM SGP observational domain.

The surface heat fluxes have similar diurnal patterns for the three daytime periods, with peak values of 145-160 W m<sup>-2</sup> for the sensible heat flux and 290-320 W m<sup>-2</sup> for the latent heat flux. ARSCL cloud occurrence frequencies show signatures of cumulus clouds each day, commonly indicating the passage of vertically developed clouds. The LCL shows a large time variation consistent with daytime cloud-base height. Cumulus clouds begin to

appear in the morning and continue until around 16 LT, during which time cloud-base heights gradually rise. ARSCL cloud top is highest on day 1 (3700-3900 m), and lower on day 2 (2000-2500 m) and day 3 (2000-2900 m); actual cloud tops may have been higher in the area since ARSCL is based on measurements of a narrow vertical column.

Figure 3 shows time-height cross sections of cloud fraction for the model domains with the LCL and level of free convection (LFC) simulated by WRF-FASTER and DHARMA PRNA. Cloud fraction at each altitude is defined as the fraction of grid boxes in which cloud condensate is larger than  $0.01 \text{ g kg}^{-1}$ . The LCL and LFC are based on air properties at the lowest level in the models. Both models produce the daytime evolution of a cumulus-topped boundary layer during the three daytime periods. The cloud fraction profiles have peak values near cloud base and decrease with height. The LCL and LFC show significant diurnal variations; during daytimes they agree well with each other and the cloud-base height; this indicates that there are positively buoyant “active cumulus clouds” in the three daytime periods (see *Stull* [1985] for dynamics-based cumulus classification, *Wilde et al.* [1985] for LCL variation and cumulus onset, and *Endo* [2009] for LFC variation and active cumulus onset). After each nightfall, WRF-FASTER reproduces the observed drop in LCL associated with development of a stable nocturnal boundary layer, whereas DHARMA systematically overestimates nocturnal LCL; in sensitivity tests shown in Part I, DHARMA reproduces the nightly drop in LCL only when using a large-scale forcing dataset generated from a WRF simulation (*Vogelmann et al.* [2015], cf. their Fig. 7); this difference is noted but not addressed here since WRF-FASTER and DHARMA LCL and cloud base are nearly identical during sequential daytimes (as shown below) and nocturnal boundary layer evolution is not a focus of this study.

Cloud fraction is very similar for the two models except that it's slightly larger in DHARMA than in WRF-FASTER. Although domain-computed cloud fraction is a measure

different from the vertically-pointing ARSCL observation, some common features can be seen in the observed and modeled clouds: the onset of boundary-layer-top cloud occurrence circa 9 LT daily, the gradual increase in cloud-base height during daytime, and the variability in vertical extent diurnally and from day to day. Both models miss the early morning clouds observed before 9 LT of day 1 and around 6 LT on day 2. These early morning clouds form before the dry convective boundary layer develops and the LCL reaches those altitudes; therefore, the clouds are likely associated with the dynamics in the nocturnal boundary layer such as shear instability and gravity waves. We found that more nocturnal clouds are generated in a sensitivity run using geostrophic forcing (not shown); however, that is beyond the scope of this study focused on daytime cumulus fields sampled by the aircraft observations.

Figure 4 compares temporal variations of the observed and simulated bulk cloud properties in terms of cloud fraction and liquid water path. Two estimates of cloud fraction are used to encompass the measurement uncertainty -- the sky cover from the total sky imager (TSI) and the maximum cloud occurrence below 5000 m in the ARSCL product. For further details on the two estimates, see *Wu et al.* [2014] and Appendix B in *Vogelmann et al.* [2015]. Liquid water path is from the MWRRet retrieval from surface-based microwave radiometer measurements [*Turner et al.*, 2007]. The observed values are hourly averages to match the domain-averaged values from the simulations. The two models agree in terms of the timing and duration of the cloudy periods and mid-day tendencies of cloud fraction and liquid water path, which decrease on days 1 and 2 and increase on day 3. As discussed earlier, the observed early morning clouds are missed by both models, which are not configured for simulating nocturnal cloud; the simulations show better agreement with observations at midday. DHARMA tends to produce greater cloud fraction and liquid water path than WRF-FASTER, which is especially apparent in the late afternoon of day 3 and as

also seen in Figure 3.

Figure 5 shows vertical profiles of cloud fraction, potential temperature, and water vapor mixing ratio for the three days. The profiles include the domain means from WRF-FASTER and DHARMA simulations and the values from the ARM VARANAL forcing dataset, both at the time of the rawinsonde launches (1130 LT; dotted lines) and at the midpoint of the 15-30-minute aircraft spiral at the end of the daily flight (1330, 1240, and 1330 LT; plus symbols). The aircraft data are averaged in 100-m height bins. Note that potential temperature from the day 1 sonde is not shown since it contains erroneous values (unrealistically high absolute instability in the lowest 2000 m). The two models commonly reproduced the observed thermodynamic structure consisting of cloud and subcloud layers. The simulated sub-cloud layer profiles are notably better mixed than the sonde-measured profiles at 1130 LT (blue and red dashed lines versus dotted line) but appear more similar in structure to the aircraft measurements 1–2 hours later (blue and red solid lines versus cross symbols). Compared to the vertical profiles from the aircraft observations and simulations, those from the ARM VARANAL forcing dataset tend to show a more stable stratification below the middle of the cloud layers. The shapes of the VARANAL vertical profiles are similar to the earlier sonde observations even at the later flight times, which is likely because the VARANAL dataset uses the sonde profiles in combination with coarse-resolution reanalysis data. In many respects, the LES-simulated profiles are closer to the in-situ-measurement than to the ARM VARANAL profiles, despite the fact that ARM VARANAL is used for relaxation in the simulations.

In Part I *Vogelmann et al.* [2015] examined the effect of relaxation on cloud bulk properties and found that thermodynamic relaxation can circumvent some of the uncertainties in the forcing datasets, particularly in reducing the drift of the simulations in the later days of the simulations. The results here suggest that, although relaxation would be effective in



correctly setting up the stratification in the morning and upper levels for the later days, weak relaxation is recommended (as in the 12-h timescale used here) to moderate the influence of uncertainties in the reference profiles on the LES-simulated thermodynamic structure, since the observed thermodynamic structure is not always well represented in VARANAL for this case.

Figure 6 shows the time-height variations of cloud water mixing ratio and droplet concentration averaged over cloudy grid points at each altitude in WRF-FASTER and DHARMA. Note that the number of averaged cloudy grid points decreases with height along with the decrease in the domain cloud fraction (Fig. 3). It is commonly seen in the two simulations that the cloud-mean cloud water (Figs. 6a and 6b) increases with height as a result of the condensation from lifting, which dominates dilution from entrainment (the height variation of cloud water will be examined further below). The time variation in the cloud water profile is associated with changes in cloud-base and cloud-top heights, where the relationship among the three does not significantly vary from day to day.

The cloud droplet number concentrations (Figs. 6c and 6d) tend to peak a few hundred meters above cloud base. Different from the cloud water mixing ratio, cloud droplet concentrations decrease from day 1 to day 3 in both the models, which could be attributed to the decrease in aerosol accumulation mode number concentration during the 60-hour period (Fig. 1). The droplet number concentrations are larger in WRF than in DHARMA PRNA, owing in part to differences in treatment of aerosol and activation, as discussed below.

The volume-mean radius (Figs. 6e and 6f) corresponds to the radius when all cloud droplets contain the same amount of liquid water. The volume-mean radius,  $r_v$ , is calculated from droplet concentration  $N_c$  and liquid water content  $LWC$ ,

$$\frac{4}{3}\pi r_v^3 \rho_w = \frac{LWC}{N_c}$$

$$r_v = \left( \frac{3 LWC}{4\pi\rho_w N_c} \right)^{1/3}$$

where  $\rho_w$  is the density of liquid water. The simulated volume-mean radius increases with height due to the increasing cloud water mixing ratio and decreasing droplet concentration, which represents the growth of cloud droplets (i.e., more liquid water per particle).

In summary, the two LES models capture general features of the observed diurnal evolution of the cumulus-topped boundary layer that repeats over three consecutive days, although some features are missed such as the early morning clouds. The two simulations produce similar cloud water distribution profiles over time, and droplet number concentration responds similarly to the specified time-varying aerosol properties.

#### 4. Cloud Microphysical Properties

In this section, we examine the microphysical properties of the simulated cumulus clouds through comparison with the in-situ aircraft measurements. In addition to the baseline simulations by WRF-FASTER and DHARMA, we also analyze the results from DHARMA for bin microphysics simulations, and sensitivity tests addressing the treatment of supersaturation, aerosol advection, and activation.

We first examine cloud properties in all of the simulations listed in Table 1. Figure 7 presents time series of total cloud fraction, liquid water path, and cloud-mean droplet number concentration for each, including the baseline simulations WRF-FASTER and DHARMA PRNA. Note that the BIN and BIN-K simulations are performed only for days 1 and 2 to examine the microphysical variability of the deeper clouds sampled on day 1 and the shallower clouds sampled on day 2 (clouds sampled on day 3 are also shallow). The six DHARMA simulations show no apparent differences in total cloud fraction and liquid water path (Figs. 7a and b); thus, the differences between WRF-FASTER and DHARMA baseline

simulations, discussed in the previous section, are likely due to different dynamics and physics other than microphysics.

In contrast to the cloud fraction and liquid water path, the droplet concentration averaged over cloudy grid cells (Fig. 7c) varies substantially among the simulations. The cloud-mean droplet concentration is almost the same in BIN and PRNA (respectively, the solid green and red lines), which is attributable to the fact that the double-moment bulk microphysics scheme used in PRNA was refined using the bin microphysics scheme as a reference in past simulations of maritime clouds. While DINA (black line) also produces droplet concentrations similar to PRNA and BIN, MAXS (orange line) shows larger values on the days 1 and 2. The MAXS and DINA and results suggest that the prediction of grid-scale supersaturation is a major difference between the WRF-FASTER and DHARMA PRNA baseline simulations, while the impact of aerosol advection is minor in this case. It is notable that WRF-FASTER and MAXS droplet concentrations match quite closely during cloud onset on days 1 and 2, consistent with the fact that they treat supersaturation most similarly; differences in the treatment of supersaturation can therefore explain the marked differences seen at the onset of day 1 in Figs. 6c and d. PRNA-K (red dashed line) and BIN-K, both of which have aerosol hygroscopicity specified to correspond to ammonium bisulfate, have much larger droplet concentrations than their respective PRNA and BIN simulations, indicating the importance of aerosol hygroscopicity to cloud droplet concentration for both the bulk and bin microphysics schemes in this case.

To evaluate the variability of simulated microphysical quantities using aircraft in-situ measurements, owing to the combination of large diurnal variability in cloud-base height and large vertical trends in microphysical quantities above cloud base, it is important to compare simulated cloud fields at the same height above cloud base as the aircraft. The temporal variation of cloud-base height is therefore first estimated from observed and simulated LCL

(calculated from near-surface air properties), which can approximate the cloud-base height of boundary layer cumulus clouds (see Fig. 3). Figure 8 shows the locations of the five surface stations closest to the flight path (ARM Central Facility [SGP], and four Oklahoma Mesonet stations [BLAC, BRUB, REDR, and BREC]) along with the time series of their calculated LCL values. The spread in values from the five surface stations is used as the observational uncertainty that reflects the regional variation in LCL. Three LCL time series (mean LCL, mean LCL plus one standard deviation, and mean LCL minus one standard deviation) are used to characterize the observational cloud-base variations. The height differences between the aircraft altitude and these three time-varying cloud-base height estimates are matched by sampling the simulations at the same height differences relative to the modeled cloud-base heights. We approximate the unbiased cloud-field sampling of the RACORO flights (i.e., sampling along a fixed flight path without deviation to pursue well-developed clouds atypical of the field) by sampling all simulated cloudy grids for a given time and relative height, which includes the variability from clouds at different stages of development and in-cloud variability as in the observed cloud field.

Following this sampling strategy, Figure 9 compares the observed and simulated LWC and droplet number concentration for the time periods and altitudes corresponding to the aircraft measurements. During RACORO, the calibration of the Cloud Aerosol and Spectrometer (CAS) was more stable than that of the Forward Scattering Spectrometer Probe (FSSP) so the total droplet concentration was obtained from the CAS. Comparison of the LWC derived from the CAS size distributions against the best reference value of the Gerber Probe showed differences of approximately 25% on average throughout RACORO. For the day 1 flight on May 22, the Gerber LWC is 38% smaller than CAS LWC on average and, for the day 2 flight on May 23, the Gerber LWC is 34% smaller than the CAS. To address the large scatter associated with the use of different probes to describe mass and number

concentrations in these clouds, the LWC is derived from the CAS size distributions but multiplied by the Gerber/CAS LWC ratio on the appropriate day; the volume-mean radius is derived from the CAS size distributions (see Appendix B for more details). For each simulation (column), the sampled cloud fields are grouped by the horizontal flight leg (row). The observed values are indicated by the black dots. The LCL sampling strategy gives three groups of points for the simulated values, indicated by colors, corresponding to the three sampling height time series using the mean (yellow) and mean plus/minus standard deviation (red/blue) of the estimated cloud-base height. Legs 1 and 2 are sampled near cloud top and Leg 4 near cloud base. Leg 3 is not shown since it sampled only a small number of cloud points.

In the observations and all simulations, number concentration varies with LWC almost linearly at the three levels. The quasi-linear relationship likely corresponds to the gradient of microphysical quantities between the cloud centers and cloud edges at given height, where maximum values are located near cloud center and the values decrease and approach zero near cloud edge. The LWC-droplet concentration line turns clockwise with height because of the increase of condensed water towards cloud top and a small decrease in number concentration with height (see Fig. 6). Differences among the three colors are most apparent at the lowest level because LWC increases most rapidly near cloud base. The red (blue) points dominate in the highest (lowest) leg because there are fewer cloudy points at the sampling heights determined using the mean plus (minus) one standard deviation of the cloud-base height.

While the value range of LWC is not very different among the simulations except for Leg 1 of BIN-K, droplet concentration varies notably among them. The range of droplet concentrations in WRF-FASTER (column 1) agree well with the observations for Legs 2 and 4 but misses the largest values near cloud top (Leg 1). DHARMA PRNA (column 2)

produces a droplet number range similar to BIN (column 4), which is slightly smaller than WRF-FASTER. The difference between the two baseline simulations (WRF-FASTER and DHARMA PRNA) is attributable to the different treatment of predicted supersaturation. This is because MAXS (column 3) emulates the treatment in WRF-FASTER and provides droplet concentrations similar to WRF-FASTER, and DINA (not shown) is similar to PRNA (e.g., see Fig. 7). In BIN-K (column 5), the large increase in hygroscopicity by changing the observation-derived value ( $\sim 0.10$ ) (column 2) to that assuming pure ammonium bisulfate ( $= 0.55$ ) significantly increases droplet concentration beyond the observed values. A similarly large impact of the increased hygroscopicity is found for the bulk microphysics scheme in PRNA (not shown). This result indicates the importance of the observation-derived hygroscopicity values to the adequate simulation of cloud droplet concentration in this case.

Given that there is the large increase in LWC with height (shown in Fig. 9), the vertical distribution of the LWC is examined in Figure 10. Shown are the height-dependent probability density functions of LWC in the simulations by WRF-FASTER, DHARMA PRNA, and BIN for 11-12 LT on day 1, which overlaps the periods of horizontal legs 1 and 2 shown in Fig. 9. The probability density functions are computed from the cloudy grids for the period as a function of height using  $0.1 \text{ g m}^{-3}$  sized bins, and are compared to the LWC for an air parcel lifted adiabatically from cloud base (red line). In all simulations shown, most values are much smaller than for the adiabatic parcel, suggesting the strong influence of entrainment of dry environmental air with cloudy air [cf. *de Rooy et al.*, 2013; *Lu et al.*, 2013a]. In the lower half of the cloud layers, the LWC maxima follow the adiabatic value. This indicates that the simulated parcels rising from cloud base maintain an adiabatic component, although the predominant values (red) are much smaller than the adiabatic value (less than  $0.5 \text{ g m}^{-3}$ ). In the upper half of the clouds, the maxima do not follow the adiabatic parcel; i.e., the clouds increasingly lose adiabaticity away from cloud base as they are

affected by entrainment. Since the observed maxima in the upper levels (Fig. 9;  $2.1 \text{ g m}^{-3}$  for Leg 1 and  $1.5 \text{ g m}^{-3}$  for Leg 2) are also less than the adiabatic parcel, this suggests that the observed cloud portions are not adiabatic due to dilution by ambient unsaturated air.

Figure 11 shows scatter plots of LWC and droplet concentration for the day 2 flight. On day 2, the cloud layer is much shallower than on day 1 (shown in Fig. 2), and all of the horizontal legs are sampled close to cloud base for these time periods. The three horizontal legs show characteristics similar to Leg 4 of the day 1 flight. The small LWC values have large differences among the three colors, meaning that there is a large rate of increase in LWC with height, which means that the uncertainty in the cloud-base-height estimation has a large impact on the analysis (e.g., for Leg 3, the difference in the maximum LWC between red and blue is  $0.5 \text{ g m}^{-3}$ ). In all simulations, the observed LWCs fall within the range of the simulated values. Similar to the day 1 flight, there is a clear impact of the supersaturation treatment between PRNA and MAXS, and from the different hygroscopicities between BIN and BIN-K. The day 3 clouds are also shallow during the flight time and have similar microphysical relationship as for day 2 (not shown).

The microphysical variability is further examined using volume-mean radius diagnosed from LWC and cloud droplet concentration. Figure 12 shows the relationship between LWC and volume-mean radius from the observations and simulations for the day 1 flight in Fig. 8. The simulated patterns are similar to the observations in that they have similar increases in the minimum volume-mean radius with LWC. The simulations with bulk microphysics tend to underestimate volume-mean radius for small LWC in Legs 1 and 2 because of larger droplet concentrations, as the differences in volume-mean radius among the models correspond to those in droplet concentration. Figure 13 is as in Fig. 12 but for the day 2 flight shown in Fig. 11. As expected, the shallow clouds have characteristics similar to the near-cloud-base portion of the deeper clouds on day 1. As seen from the different colors,

there are large variations with height near cloud base; however, the observed values of the volume-mean radius are almost within the range of the simulated values.

The  $r_v$ -LWC relationship is often used to examine mixing state of the clouds and underlying entrainment-mixing scenario [e.g., *Gerber et al.*, 2008; *Lehmann et al.*, 2009; *Lu et al.*, 2011], i.e., how cloud droplets are influenced by subsaturated entrained air via the evaporation and turbulent mixing. When the turbulent mixing is much faster than the evaporation, all droplets face the entrained subsaturated air and reduce its radius; therefore  $r_v$  decreases with decreasing LWC (homogeneous mixing). On the other hand, when the evaporation is much faster than turbulent mixing, some droplets face inhomogeneously distributed subsaturated air and evaporate completely; therefore  $r_v$  doesn't change with decreasing LWC because of decreasing number concentration (inhomogeneous mixing). In the simulations here we are not using a parameterization of subgrid-scale entrainment mixing, such as suggested by *Lu et al.* [2013b]; thus, the microphysical variability shown here is only from resolved grid-scale entrainment and mixing (i.e., subgrid-scale mixing is homogeneous). However, the simulated  $r_v$ -LWC relationship shows the signatures of both homogeneous mixing and inhomogeneous mixing as in the observations. It thus appears that the entrainment mixing process is partly resolved at the LES grid scale in this case. More quantitative analysis would be necessary to explore the impacts of grid size and subgrid-scale mixing parameterization.

Vertical motion is analyzed in Figure 14, which shows the relationship between LWC and vertical motion from the aircraft observations and simulations for the day 1 flight in Fig. 8. The simulated distribution of LWC and vertical motion largely overlaps the observed distribution, except that the maximum vertical motion is slightly underestimated in Leg 1 in some simulations ( $13 \text{ m s}^{-1}$  in the observation;  $10 \text{ m s}^{-1}$  in WRF-FASTER;  $8 \text{ m s}^{-1}$  in MAXS and BIN). The vertical motion concentrates within a small range in Leg 4 near the



cloud base, and spans a wide range in Legs 1 and 2 near the cloud top. Large vertical motion tends to occur with large LWC, as it would be associated with the air mass in the vicinity of the cloud center that experienced substantial buoyancy production from condensation and less dilution from entrainment.

Figure 15 is as in Fig. 14 but for the day 2 flight shown in Fig. 11. Similar to other microphysical aspects discussed, the vertical motion also has the characteristics found near the cloud base of the deeper clouds on day 1. It is commonly seen in all legs that the vertical motion spans the small range similar in Leg 4 of the day 1 flight (Fig. 14). Larger vertical motion again tends to be associated with larger LWC. Different colors (sampling heights) show similar ranges of vertical motion in Legs 1 and 2, which suggests that the air parcel near cloud base is still associated with forced convection from the subcloud layer and is not completely driven by positive buoyancy caused by the condensation.

## **5. Summary and Discussion**

A set of LES runs is performed for continental boundary layer cumulus clouds under temporally-varying large-scale and aerosol conditions during the RACORO field campaign at the ARM SGP site. Comprehensive evaluations of the LES use in-situ aircraft and ground-based observations of cloud macrophysical and microphysical properties.

The comparison shows that both LES models capture the observed time evolution of the cumulus-topped boundary layers driven by the surface sensible and latent heat fluxes over three daytime periods, with the exception of some early morning clouds observed before the development of the convective boundary layer on the first day. The models agree well on the onset and end time of the cumulus clouds, and on the thermodynamic structure in terms of the vertical profiles of potential temperature and water vapor. The simulations produce thermodynamic structures similar to the aircraft profiling observations. The difference in

thermodynamic profiles between the simulations and the aircraft observations is smaller than the difference from the reference value given in the forcing dataset, suggesting that the models faithfully simulate the boundary layer dynamics in a manner that is not well represented in the large-scale forcing dataset.

In both models, the cloud-mean cloud water mixing ratio increases with height as a result of condensation from lifting, which dominates dilution from entrainment, and varies with relative height from cloud base and with changes in cloud base height. The cloud-mean cloud droplet concentrations decrease day to day in both the models, which is attributed to the decrease in aerosol accumulation mode number concentration during the 60-hour period. The cloud-mean volume-mean radius diagnosed from cloud water mixing ratio and droplet concentration also shows the increase with height, meaning that the growth of cloud droplets are represented in the simulations. The results indicate that, with the temporally-varying forcing and aerosol size distribution profiles, the models can generate the complex microphysical variations tied to the varying cloud fields.

Besides the cloud-mean values, the simulated microphysical variability is evaluated by the aircraft observations for the horizontal legs of the flight triangles. To compare with the aircraft observations, the simulated cloud fields were sampled at a time-dependent relative height above cloud base to match that of the aircraft track in the evolving cloud-topped boundary layer. This approach considers time variations of both the aircraft altitudes and evolving cloud fields. The models commonly simulate LWC in a linear relationship with droplet number concentration. The LWC-droplet concentration line turns clockwise with height because of the large increase of condensed water toward cloud top and the decrease in droplet number concentration with height.

Based on the height variation of a probability distribution function of LWC and that computed by an adiabatic parcel model, the evolution of LWC is strongly influenced by

entrainment.

Compared with the aircraft measurements, the range of simulated droplet concentrations is underestimated in the leg at the highest altitude, but differences remain on par with the potential estimation errors in the measurements; in other legs, the observed droplet concentration falls within the range of simulated values. The range of droplet concentrations in WRF-FASTER is generally larger than in DHARMA; concentrations are most similar when DHARMA adopts the WRF-FASTER treatment of supersaturation. The range of simulated LWC values is not very different between the models, and the observed LWC spans the range of simulated values. The models also captured some of the observed features in  $r_v$ -LWC and  $w$ -LWC distributions. The results suggest that, given appropriate constraints on aerosol properties, both LES models have the ability to produce cloud microphysical variability comparable to the aircraft measurements given estimation errors as suggested by the different estimations of LWC and number concentration from the different probes.

Additional sensitivity experiments using DHARMA show that: 1) prognostic treatment of aerosol number concentration has only a minor effect on cloud droplet number concentrations in this case; 2) prediction of supersaturation has a strong influence on cloud droplet number concentration, which is at least one cause of the different number concentrations between the baseline simulations by WRF-FASTER and DHARMA PRNA; and 3) aerosol hygroscopicity significantly affects cloud droplet number concentration but has little effect on liquid water path and cloud fraction, presumably because collision-coalescence is not a strong process in this case and thus cloud microphysics has little impact on cloud dynamics here.

This case study offers a benchmark of LES performance in simulating continental cumulus clouds under varying forcing and aerosol conditions, in which the models are

constrained by unique trimodal aerosol size distribution profiles and hygroscopicity parameters and are evaluated by inter-model comparison and comprehensive aircraft and ground-based observations. The developed case study can contribute to linking extensive observations to the improvement of climate model parameterizations. Such an application is reported in Part III of this series [Lin *et al.*, 2015], which assesses deficiencies in the Community Atmosphere Model-5 (CAM5) parameterizations. The study uses the LES output presented, together with observations and simulations with the CAM5 single-column model driven by the same large-scale and surface forcings used here. Recent enhancement of observations (e.g., implementation of a new set of scanning cloud radars and land-surface measurements at the ARM sites [Mather and Voyles, 2013]) will allow for more comprehensive model validation and further utilization of LES to study boundary layer clouds and their interaction with the underlying land surface.

## Acknowledgements

This research was supported by the U.S. Department of Energy Science Office of Biological and Environmental Research Program under the Earth System Modeling Program via the FASTER Project (<http://www.bnl.gov/faster/>), and the Atmospheric System Research Program via DE-SC00112704. Observational data sets were obtained from the U.S.

Department of Energy

ARM Climate Research Facility (<https://www.arm.gov/>) and Oklahoma Mesonet

(<https://www.mesonet.org/>). Processed forcing and observational data sets used for the

three-day case study are also aggregated as an ARM PI data product

([http://iop.archive.arm.gov/arm-iop/0pi-data/vogelmann/racoro/case\\_studies](http://iop.archive.arm.gov/arm-iop/0pi-data/vogelmann/racoro/case_studies)). The research

utilized resources at the New York Center for Computational Sciences. Authors appreciate

Prof. Marat Khairoutdinov for his helpful comments on LES configurations, Dr. Kwinten

Van Weverberg and Dr. Hugh Morrison for providing and helping to use the two-moment

microphysics scheme and Dr. Peter Blossey for providing the interface to RRTM radiation scheme in WRF-FASTER.

## **Appendix A: Bin microphysics in DHARMA**

DHARMA implements a modified version of the CARMA size-resolved microphysics model [Ackerman *et al.*, 1995; Jensen *et al.*, 1998], in which cloud, drizzle and rain drops are represented as a continuous size distribution that is discretized into 36 size bins. Aerosol activation introduces droplets into the smallest bin, which corresponds to a droplet diameter of 2  $\mu\text{m}$ . The mean mass of each successive bin is twice that of the preceding bin.

Hydrometeor fall speeds and collision rates are calculated following *Böhm* [1989, 1992a, 1992b, 1999, 2004], wherein the mean maximum dimension and aspect ratio of large drop deformation follows *Green* [1975].

For the simulations in this study, the total size distribution is discretized into 36 bins (each a prognostic variable), wherein the smallest mass bin corresponds to an aerosol particle with a dry diameter of 10 nm, given a baseline ammonium bisulfate aerosol mass density of 1.78  $\text{g cm}^{-3}$ , and the mean mass of each successive bin is 1.5 times that of the preceding bin. The hygroscopicity parameter calculated for ammonium bisulfate is generally replaced with observation-based values [Vogelmann *et al.*, 2015]. The droplets smaller than 25  $\mu\text{m}$  radius are considered as cloud droplets and compared to the aircraft observations and bulk microphysics scheme.

## **Appendix B: Uncertainty analysis and post-measurement correction of liquid water content from the aircraft observations**

There are uncertainties in microphysical quantities derived from the aircraft observations that are used for the model validations. *McFarquhar and Jackson* [2014] investigated the differences in droplet size distributions measured by Cloud and Aerosol Spectrometer (CAS) and Forward Scattering Spectrometer Probe (FSSP) during the RACORO field campaign. They compared liquid water content (LWC) derived from CAS and FSSP size distributions against bulk LWC measured by the widely-used Particle Volume Monitor (Gerber Probe). In their analysis, CAS LWC was about 25% larger than Gerber LWC, and FSSP LWC was about 15% smaller on average; however, there was also a lot of scatter in the relations for different days. They also noted that calibration of CAS was steadier than that of FSSP during RACORO, as the later had a tendency to drift and required frequent disassembly, cleaning and reassembly, and a recalibration after each flight. See *McFarquhar and Jackson* [2014] for more details.

Figure B1 compares LWC from the CAS, FSSP, and Gerber Probe for the day 1 and day 2 flights in this case study. For these particular days, different from the relationships noted above, CAS LWC is smaller than Gerber LWC, and FSSP LWC is larger than Gerber LWC. Given this result and a potential of measurement drift of FSSP, the CAS was used to determine LWC, after applying a correction factor to account for this offset from the Gerber LWC:  $LWC = CAS\ LWC / 0.6177$  for the day 1 flight and  $LWC = CAS\ LWC / 0.65578$  for the day 2 flight. It should also be noted that there are also uncertainties in the LWC measured by the Gerber Probe: past studies have estimated uncertainties on the order of 10% [*Gerber et al.*, 1994], with uncertainties increasing when larger droplets are present [e.g., *Gerber et al.*, 1999].

The droplet number concentrations ( $N_c$ ) measured by the CAS and FSSP were also

examined. According to *McFarquhar and Jackson* [2014], they were generally in good agreement during RACORO. For the flights examined here, on average the CAS  $N_c$  was 10% larger than the FSSP  $N_c$  for the day1 flight, and was 1% larger for the day 2 flight (not shown); at a specific point in time, the two probes could differ in  $N_c$  by about 50% (CAS  $N_c \sim 2 \times$  FSSP  $N_c$ ). Because the CAS calibration was more stable than that of the FSSP, the  $N_c$  from the CAS is used for the model evaluations presented here.

## References

- Abdul-Razzak, H., S. J. Ghan, and C. Rivera-Carpio (1998), A parameterization of aerosol activation: 1. Single aerosol type, *J. Geophys. Res.*, 103(D6), 6123–6131, doi:10.1029/97JD03735.
- Abdul-Razzak, H., and S. J. Ghan (2000), A parameterization of aerosol activation: 2. Multiple aerosol types, *J. Geophys. Res.*, 105 (D5), 6837–6844, doi:10.1029/1999JD901161.
- Ackerman, A. S., and Coauthors (2009), Large-eddy simulations of a drizzling, stratocumulus-topped marine boundary layer, *Mon. Wea. Rev.*, 137, 1083–1110.
- Ackerman, A.S., O.B. Toon, D.E. Stevens, A.J. Heymsfield, V. Ramanathan, and E.J. Welton (2000), Reduction of tropical cloudiness by soot, *Science*, 288, 1042–1047, doi:10.1126/science.288.5468.1042.
- Ackerman, A. S., O. B. Toon, and P. V. Hobbs (1995), A model for particle microphysics, turbulent mixing, and radiative transfer in the stratocumulus-topped marine boundary layer and comparisons with measurements, *J. Atmos. Sci.*, 52, 1204–1236.
- Blossey, P. N., C. S. Bretherton, M. Zhang, A. Cheng, S. Endo, T. Heus, Y. Liu, A. P. Lock, S. R. de Roode, and K.-M. Xu (2013), Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison, *J. Adv. Model. Earth Syst.*, 5, 234–258,

doi:10.1002/jame.20025.

- Böhm, H. (1989), A general equation for the terminal fall speed of solid hydrometeors, *J. Atmos. Sci.*, 46, 2419–2427.
- Böhm, J. P. (1992a), A general hydrodynamic theory for mixed-phase microphysics. Part I: Drag and fall speed of hydrometeors, *Atmos. Res.*, 27, 253–274.
- Böhm, J. P. (1992b), A general hydrodynamic theory for mixed-phase microphysics. Part II: Collision kernels for coalescence, *Atmos. Res.*, 27, 275–290.
- Böhm, J. P. (1999), Revision and clarification of “A general hydrodynamic theory for mixed-phase microphysics”, *Atmos. Res.*, 52, 167–176.
- Böhm, J. P. (2004), Reply to Comment on “Revision and clarification of ‘A general hydrodynamic theory for mixed-phase microphysics’ [Bohm J.P., 1999, *Atmos. Res.* 52, 167-176].”, *Atmos. Res.*, 69, 289–293, doi:10.1016/j.atmosres.2003.10.001.
- Bretherton, C. S., Krueger, S. K., Wyant, M. C., Bechtold, P., Van Meijgaard, E., Stevens, B., and Teixeira, J. (1999), A GCSS boundary-layer cloud model intercomparison study of the first ASTEX Lagrangian experiment, *Bound.-Layer Meteor.*, 93(3), 341-380.
- Brown, A. R., and Coauthors (2002), Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land, *Quart. J. Roy. Meteor. Soc.*, 128, 1075–1093.
- Browning, K. A., and Coauthors (1993), The GEWEX Cloud System Study (GCSS). *Bull. Amer. Meteor. Soc.*, 74, 387–399.
- Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, B. E. Martner (2000), Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites, *J. Appl. Meteor.*, 39, 645–665.
- Deardorff, J. W. (1980), Stratocumulus-capped mixed layers derived from a three-dimensional model, *Bound.-Layer Meteor.*, 18, 495–527.



- de Rooy, W. C., P. Bechtold, K. Fröhlich, C. Hohenegger, H. Jonker, D. Mironov, A. Pier Siebesma, A., J. Teixeira, and J.-I. Yano (2013), Entrainment and detrainment in cumulus convection: an overview, *Q. J. R. Meteorol. Soc.*, 139, 1–19, doi: 10.1002/qj.1959.
- Endo, S. (2009), Study on convective boundary layer and cumulus clouds over humid terrestrial area, *Ph.D thesis*, Nagoya University, Japan.
- Fridlind, A. M., and Coauthors (2012), A comparison of TWP-ICE observational data with cloud-resolving model results, *J. Geophys. Res.*, 117, D05204, doi:10.1029/2011JD016595.
- Gerber, H., B.G. Ahrens, and A.S. Ackerman (1994), New microphysics sensor for aircraft use, *Atmos. Res.*, 31, 235-252.
- Gerber, H. E., G. M. Frick, J. B. Jensen, and J. G. Hudson (2008), Entrainment, mixing, and microphysics in trade-wind cumulus, *J. Meteorol. Soc. Japan*, 86, 87–106.
- Gerber, H., G. Frick, and A.R. Rodi (1999), Ground-based FSSP and PVM measurements of liquid water content, *J. Atmos. Ocean. Tech.*, 16, 1143-1149.
- Ghan, S. J. and Coauthors (2011), Droplet nucleation: Physically-based parameterizations and comparative evaluation, *J. Adv. Model. Earth Syst.*, 3, M10001, doi:10.1029/2011MS000074.
- Green, A. W. (1975), An approximation for the shapes of large raindrops. *J. Appl. Meteorol.*, 14, 1578–1583.
- Jensen, E. J. and Coauthors (1998), Ice nucleation processes in upper tropospheric wave-clouds observed during SUCCESS, *Geophys. Res. Lett.*, 25, 1363–1366.
- Khairoutdinov, M. and D. Randall (2006), High-resolution simulation of shallow-to-deep convection transition over land, *J. Atmos. Sci.*, 63, 3421–3436.
- Kirkpatrick, M. P., A. S. Ackerman, D. E. Stevens, and N. N. Mansour (2006), On the application of the dynamic Smagorinsky model to large-eddy simulations of the

cloud-topped atmospheric boundary layer, *J. Atmos. Sci.*, 63, 526–546.

Lehmann, K., H. Siebert, and R. A. Shaw (2009), Homogeneous and inhomogeneous mixing in cumulus clouds: Dependence on local turbulence structure, *J. Atmos. Sci.*, 66, 3641–3659, doi:10.1175/2009JAS3012.1.

Lin, W., and Coauthors (2015), RACORO Continental Boundary Layer Cloud Investigations.

Part III: Separation of parameterization biases in Single-Column Model CAM5

simulations of shallow cumulus, *J. Geophys. Res.*, (this issue).

Lu, C., Y. Liu, and S. Niu (2011), Examination of turbulent entrainment-mixing mechanisms using a combined approach, *J. Geophys. Res.*, 116, D20207, doi:10.1029/2011JD015944.

Lu, C., Y. Liu, S. Niu, and A. Vogelmann (2013a), Empirical relationship between entrainment rate and microphysics in cumulus clouds, *Geophys. Res. Lett.*, 40, 1–6, doi:10.1002/grl.50445.

Lu, C., Y. Liu, S. Niu, S. Krueger, and T. Wagner (2013b), Exploring parameterization for turbulent entrainment-mixing processes in clouds, *J. Geophys. Res.*, 118(1), 185–194.

Mather, J. H., and J. W. Voyles (2013), The Arm Climate Research Facility: A Review of Structure and Capabilities, *Bull. Amer. Meteor. Soc.*, 94, 377–392. doi: <http://dx.doi.org/10.1175/BAMS-D-11-00218.1>.

McFarquhar, M. G. and R. Jackson, (2014), RACORO Value Added Product (VAP) microphysics data set version 1.0, [http://iop.archive.arm.gov/arm-iop/0pi-data/mcfarquhar/RACORO/RACORO\\_VAPv3.docx](http://iop.archive.arm.gov/arm-iop/0pi-data/mcfarquhar/RACORO/RACORO_VAPv3.docx), 6 pp.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.*, 102(D14), 16663–16682, doi:10.1029/97JD00237.

Morrison, H., J. Curry, and V. Khvorostyanov (2005), A new double-moment microphysics

- parameterization for application in cloud and climate models. Part I: Description, *J. Atmos. Sci.*, 62, 1665–1677.
- Morrison, H., and W. W. Grabowski (2008), Modeling supersaturation and subgrid-scale mixing with two-moment bulk warm microphysics, *J. Atmos. Sci.*, 65, 792–812, doi:10.1175/2007JAS2374.1.
- Petters, M. D., and S. M. Kreidenweis (2007), A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, *Atmos. Chem. Phys.*, 7, 1961–1971, doi: 10.5194/acp-7-1961-2007.
- Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski (2003a), Breaking the cloud parameterization deadlock, *Bull. Amer. Meteor. Soc.*, 84, 1547–1564.
- Randall, D., and Coauthors (2003b), Confronting models with data: The GEWEX Cloud System Study, *Bull. Amer. Meteor. Soc.*, 84, 455–469.
- Siebesma, A. P., and Coauthors (2003), A large-eddy simulation intercomparison study of shallow cumulus convection, *J. Atmos. Sci.*, 60, 1201–1219.
- Skamarock, W. C., and Coauthors (2008), A description of the Advanced Research WRF version 3, NCAR Tech. Note NCAR/TN-475+STR, 113 pp.
- Stevens, B., and Coauthors (2001), Simulations of trade wind cumuli under a strong inversion, *J. Atmos. Sci.*, 58, 1870–1891.
- Stevens, D. E., A. S. Ackerman, and C. S. Bretherton (2002), Effects of domain size and numerical resolution on the simulation of shallow cumulus convection, *J. Atmos. Sci.*, 59, 3285–3301.
- Stevens, B., and Coauthors (2005), Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus, *Mon. Wea. Rev.*, 133, 1443–1462.
- Stokes, G. M. and S. E. Schwartz (1994), The Atmospheric Radiation Measurement (ARM) program: Programmatic background and design of the cloud and radiation test bed, *Bull.*

*Amer. Meteor. Soc.*, 75, 1201–1221.

Stull, R. B. (1985), A fair-weather cumulus cloud classification scheme for mixed-layer studies, *J. Climate App. Meteor.*, 24, 49–56.

Toon, O. B., C. P. McKay, T. P. Ackerman, and K. Santhanam (1989), Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres, *J. Geophys. Res.*, 94(D13), 16287–16301, doi:10.1029/JD094iD13p16287.

Turner, D. D., S. A. Clough, J.C. Liljegren, E.E. Clothiaux, K. Cady-Pereira, and K. L. Gaustad (2007), Retrieving liquid water path and precipitable water vapor from Atmospheric Radiation Measurement (ARM) microwave radiometers, *IEEE Trans. Geosci. Remote Sens.*, 45, 3680–3690, doi:10.1109/TGRS.2007.903703.

Vogelmann, A. M., and Coauthors (2012), Racoro extended-term aircraft observations of boundary layer clouds, *Bull. Amer. Meteor. Soc.*, 93, 861–878.  
doi:10.1175/BAMS-D-11-00189.1

Vogelmann, A. M., and Coauthors (2015), RACORO Continental Boundary Layer Cloud Investigations. Part I: Case study generation and ensemble large-scale forcings, *J. Geophys. Res.*, (this issue).

Wilde, N. P., R. B. Stull and E. W. Eloranta, (1985), The LCL zone and cumulus onset, *J. Clim. Appl. Meteor.*, 24, 640–657.

Wu, W., Liu, Y., Jensen, M. P., Toto, T., Foster, M. J., and Long, C. N. (2014), A comparison of multiscale variations of decade-long cloud fractions from six different platforms over the Southern Great Plains in the United States, *J. Geophys. Res.* 119, 3438–3459, doi:10.1002/2013JD019813.

vanZanten, M. C., and Coauthors (2011), Controls on precipitation and cloudiness in simulations of trade - wind cumulus as observed during RICO, *J. Adv. Model. Earth Syst.*, 3, M06001, doi:10.1029/2011MS000056.

Xie, S., R. T. Cederwall, M. Zhang (2004), Developing long-term single-column model/cloud system-resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations, *J. Geophys. Res.*, 109, D011004.

Zhang, M. H., J. L. Lin, R. T. Cederwall, J. J. Yio, and S. C. Xie (2001), Objective analysis of ARM IOP data: Method and sensitivity, *Mon. Wea. Rev.*, 129, 295–311.

Zhang, M, and Coauthors (2013), CGILS: First results from an international project to understand the physical mechanisms of low cloud feedbacks in general circulation models, *J. Adv. Model. Earth Syst.*, 5, 826–842, doi:10.1002/2013MS000246.

Zhu, P. and B. Albrecht (2003), Large eddy simulation of continental shallow cumulus convection, *J. Geophys. Res.*, 108(D15), 4453.

Zhu, P., B. A. Albrecht, V. P. Ghate, Z.-D., Zhu (2010), Multiple scale simulations of stratocumulus clouds, *J. Geophys. Res.*, 115, D23201, doi:10.1029/2010JD014400.

Table 1: List of models/simulations. See text for other differences between the models.

| Model/Simulation | Microphysics              | Grid-scale supersaturation | Aerosol concentration | Hygroscopicity  |
|------------------|---------------------------|----------------------------|-----------------------|-----------------|
| WRF-FASTER       | Two-moment bulk scheme*   | Diagnostic                 | Diagnostic            | Varying (~0.1)  |
| DHARMA PRNA      | Two-moment bulk scheme*   | Prognostic                 | Prognostic            | Varying (~0.1)  |
| DHARMA DINA      | Two-moment bulk scheme    | Prognostic                 | Diagnostic            | Varying (~0.1)  |
| DHARMA MAXS      | Two-moment bulk scheme    | Diagnostic                 | Diagnostic            | Varying (~0.1)  |
| DHARMA BIN       | Size-resolving bin scheme | Prognostic                 | Prognostic            | Varying (~0.1)  |
| DHARMA PRNA-K    | Two-moment bulk scheme    | Prognostic                 | Prognostic            | Constant (0.55) |
| DHARMA BIN-K     | Size-resolving bin scheme | Prognostic                 | Prognostic            | Constant (0.55) |

\* Both WRF-FASTER and DHARMA use the Morrison et al. [2005] scheme that is differently modified in each model.

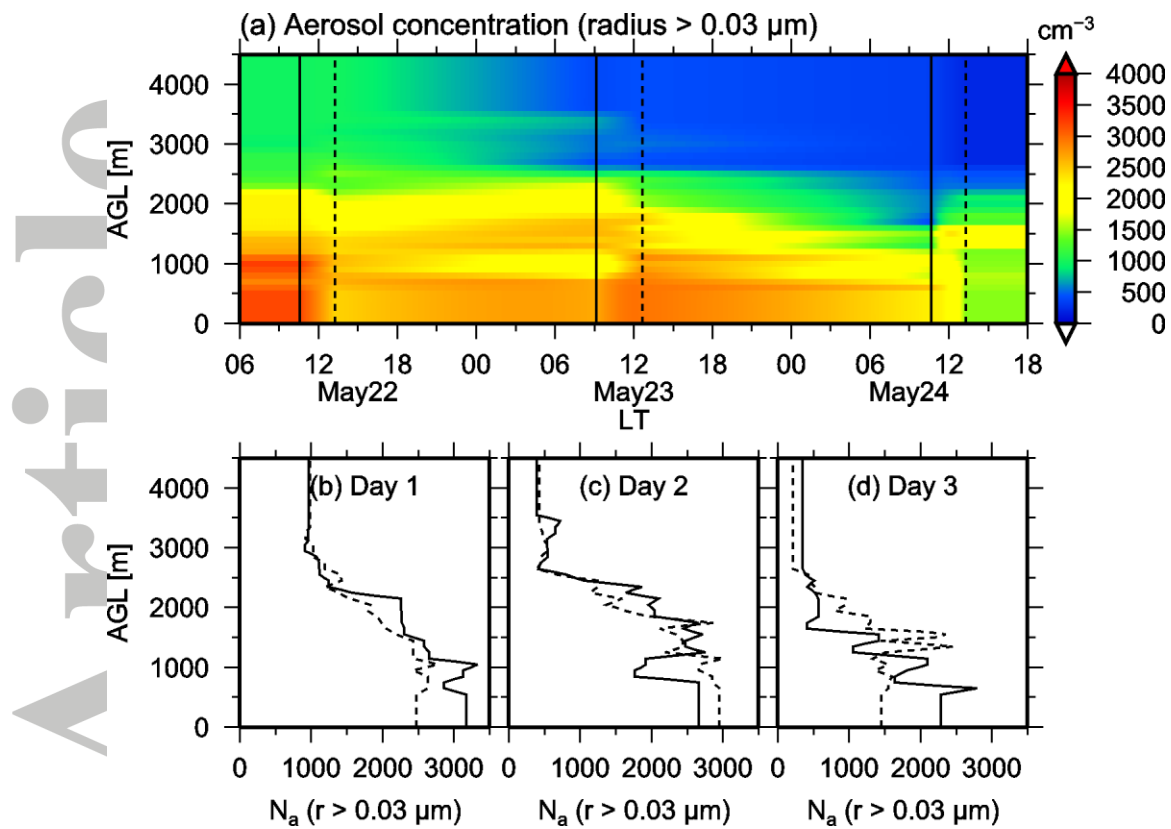


Figure 1: Time-height cross section (a) and vertical profiles (b)-(d) of number concentration of aerosol particles larger than  $0.03 \mu\text{m}$  radius. Solid and dashed lines indicate the times of the airborne measurements for each day in (a) and vertical profiles at the measurement times in (b)-(d). Time is in local time (LT).

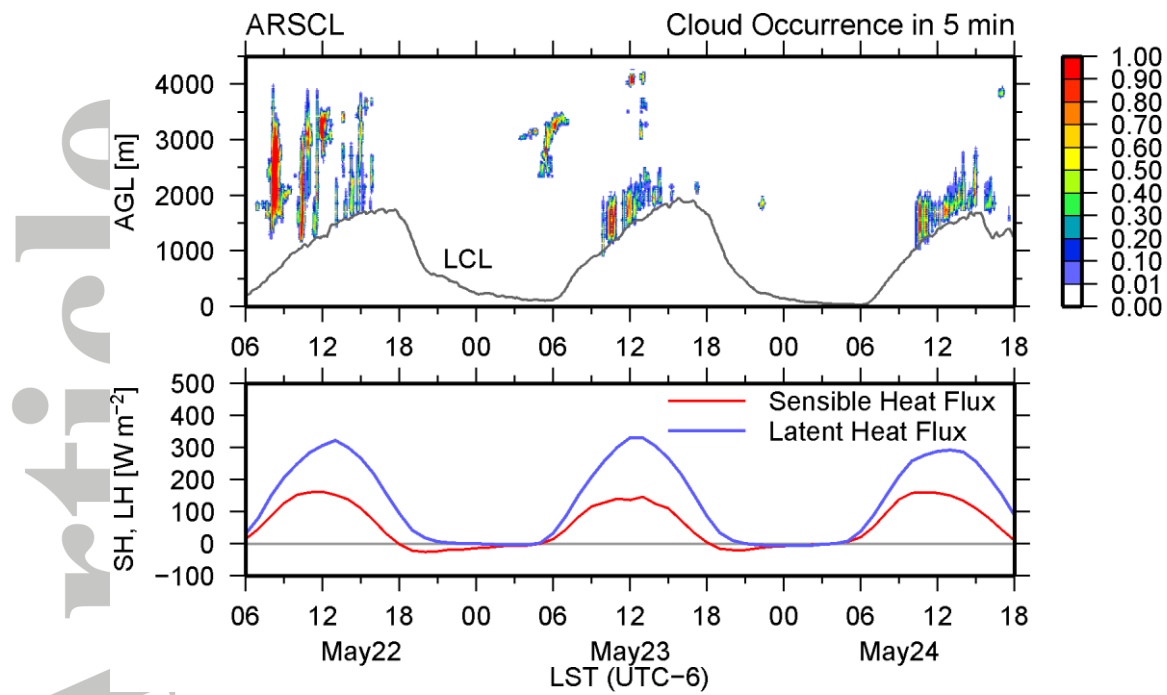


Figure 2: Time-height cross section of ARSCL cloud occurrence with LCL (gray line), and time series of the observation-based surface sensible heat flux (red line) and latent heat flux (blue line) prescribed in the simulations.



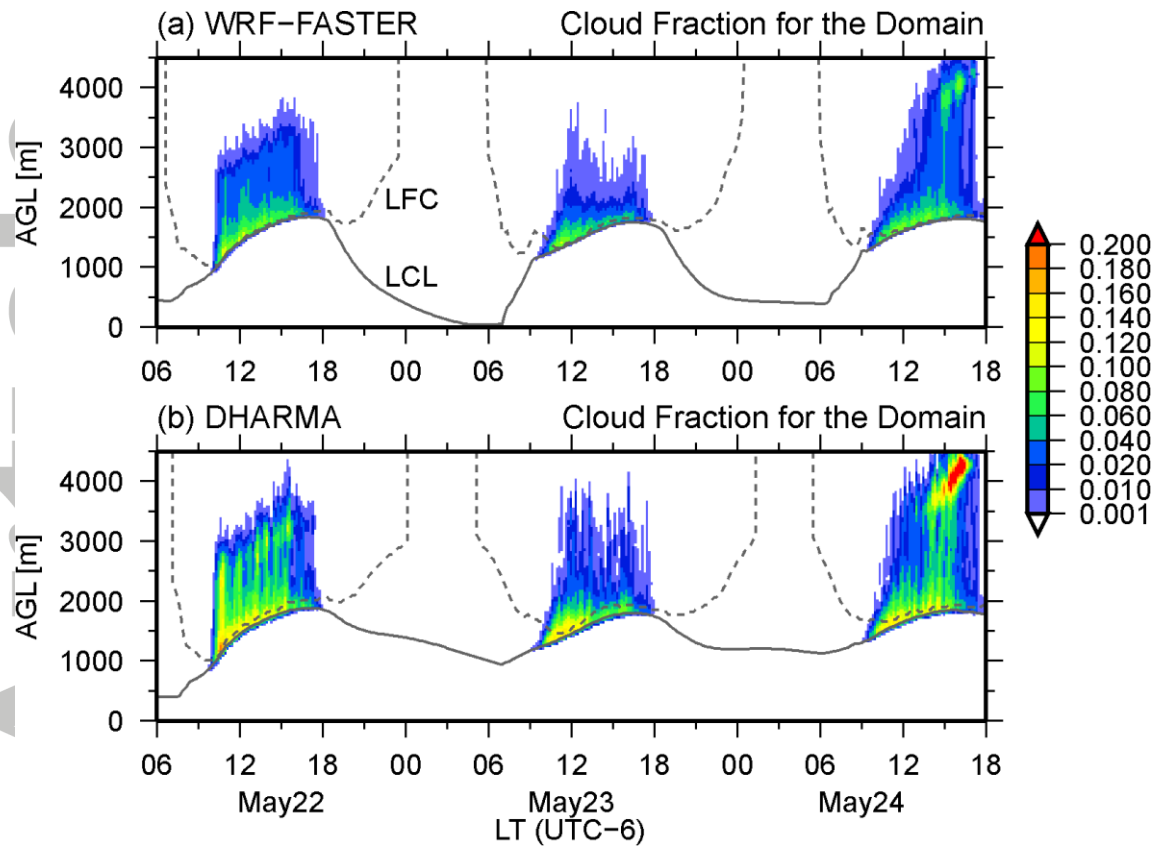


Figure 3: Time-height cross sections of spatial cloud fraction calculated from the simulations by WRF-FASTER (a) and DHARMA PRNA (b). Gray solid and dashed lines indicate LCL and LFC, respectively.

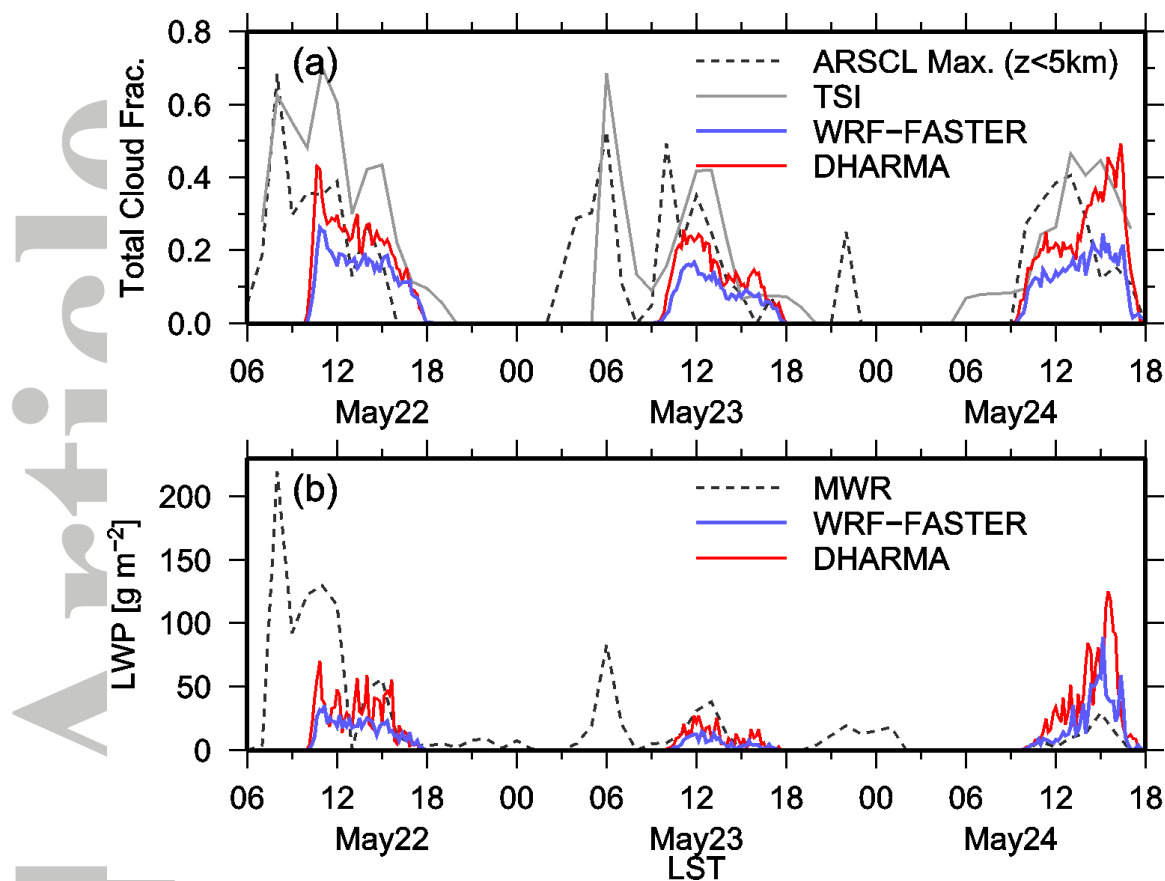


Figure 4: Time series of total cloud fraction (a) and liquid water path (b) from observations (dashed black and gray lines), and from simulations by WRF-FASTER (blue lines) and DHARMA PRNA (red lines). Dashed lines indicate the maximum cloud fraction in the ARSCL observations for altitudes below 5 km (a) and liquid water path retrieved by the microwave radiometer (b). The gray line indicates cloud fraction from the TSI.

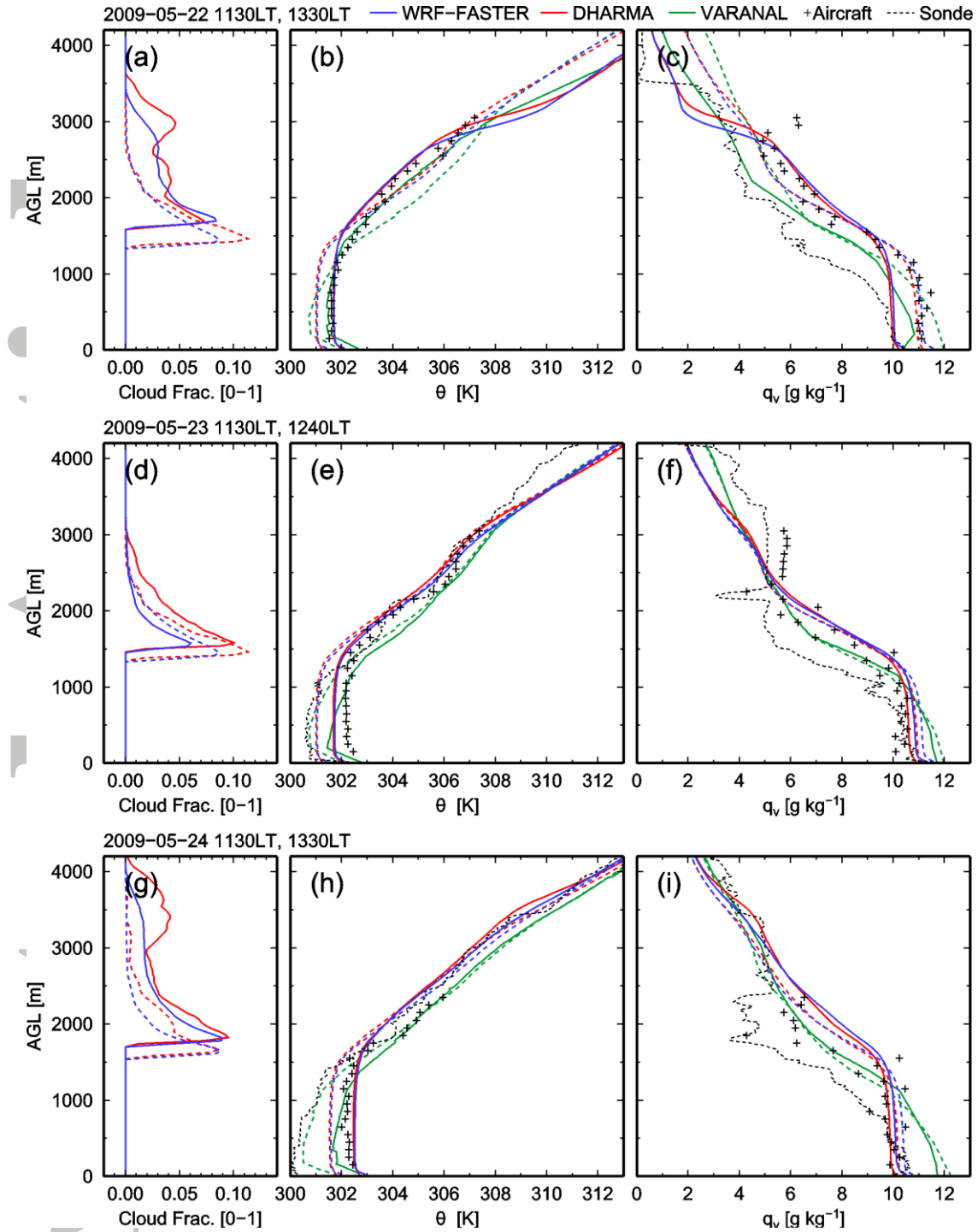


Figure 5: Vertical profiles of cloud fraction (left), potential temperature (center), and water vapor mixing ratio (right) near the time of the rawinsonde (1130LT; black dotted lines) and aircraft profiling observations (plus symbols) for day 1 (top), day 2 (middle), and day 3 (bottom). Blue, red, and green dashed (solid) lines indicate the WRF-FASTER and DHARMA PRNA simulations and reference values from the ARM VARANAL forcing at around the time of the sondes (aircraft profiles). Potential temperature from sonde is not shown for day 1 (b) since it has erroneous values.

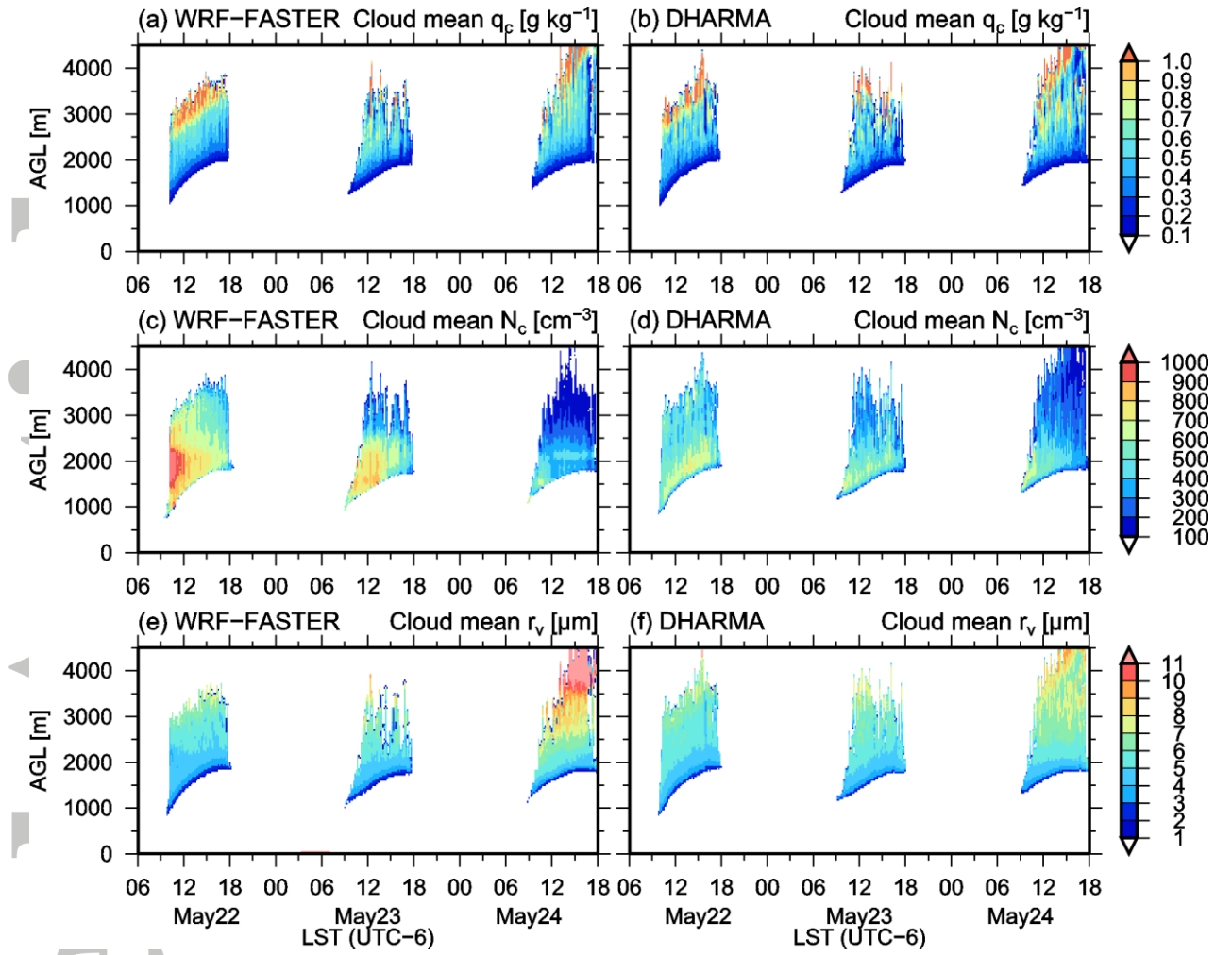


Figure 6: Time-height cross sections of cloud water mixing ratio (a, b), cloud droplet number concentration (c, d), and volume mean radius in the simulations by WRF-FASTER (left) and DHARMA PRNA (right). Values are averaged over cloudy grid points that have cloud water mixing ratios larger than  $10^{-2} \text{ g kg}^{-1}$ .

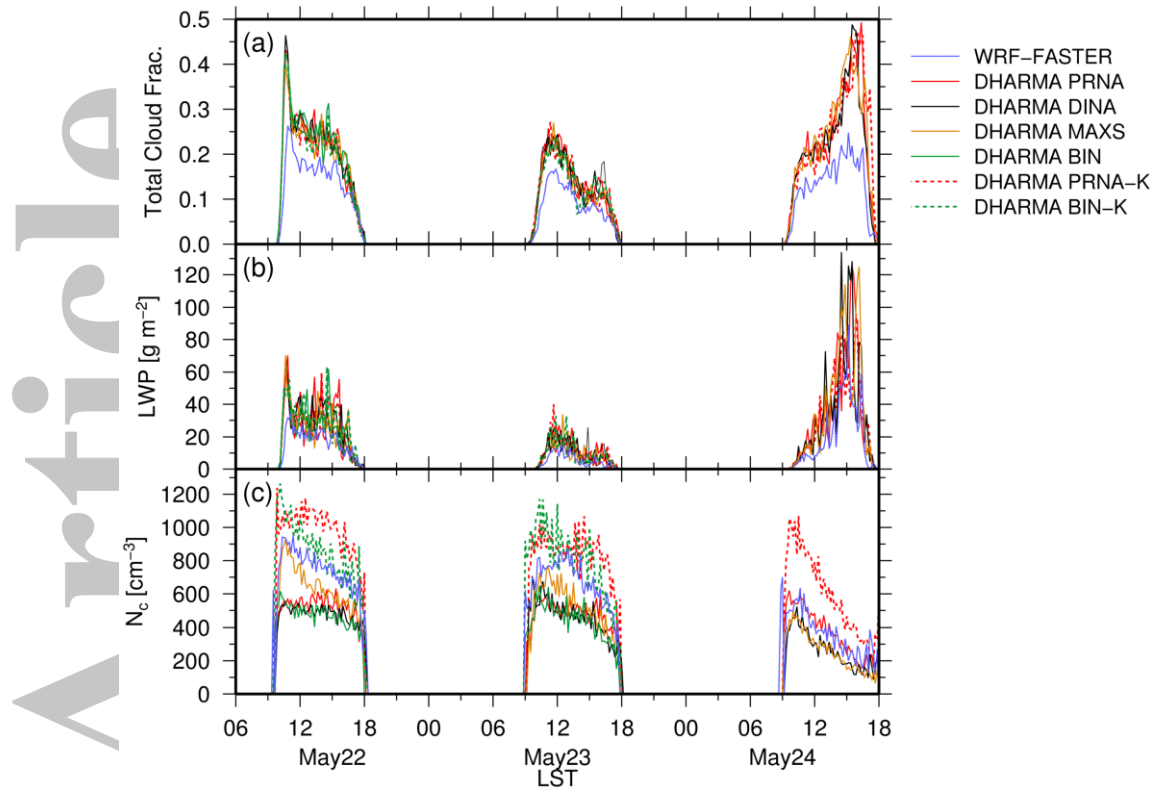


Figure 7: Time series of total cloud fraction (a), liquid water path (b), and cloud-mean droplet number concentration (c) from the simulations by WRF-FASTER and DHARMA. See text for the details of the configurations of the sensitivity studies.

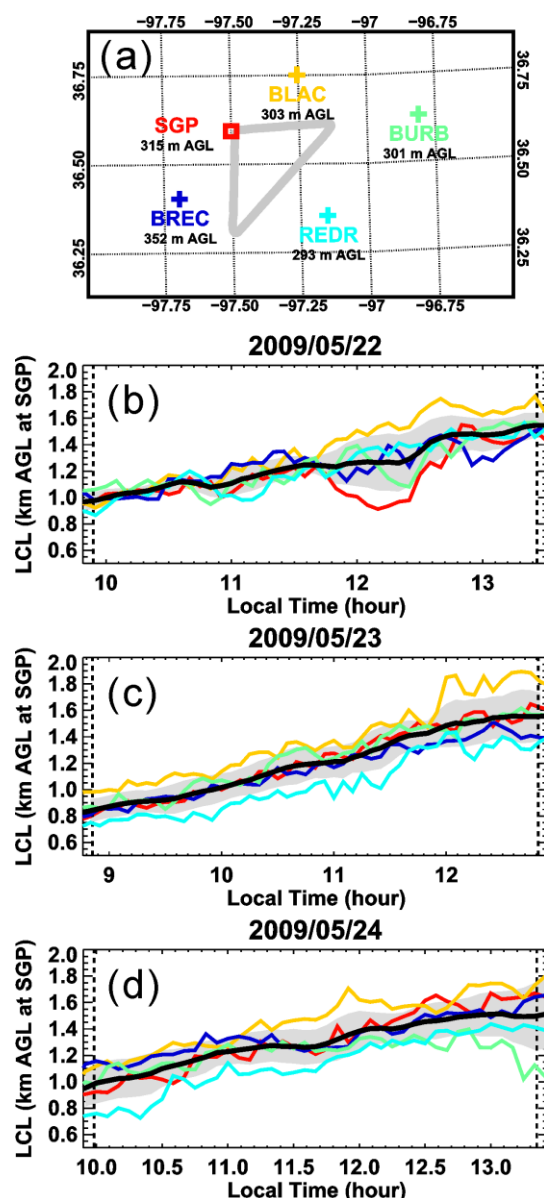


Figure 8: (a) Locations of the horizontal legs of the triangular aircraft flight pattern over which observations were obtained (gray line), and five adjacent surface stations (ARM SGP Central Facility and four Oklahoma Mesonet stations). (b - d) Time series of LCL estimated from air properties observed at the five surface stations for each flight period. Colors correspond to the surface stations in (a). The black line and gray shading indicate, respectively, the mean and standard deviations of the surface station values.

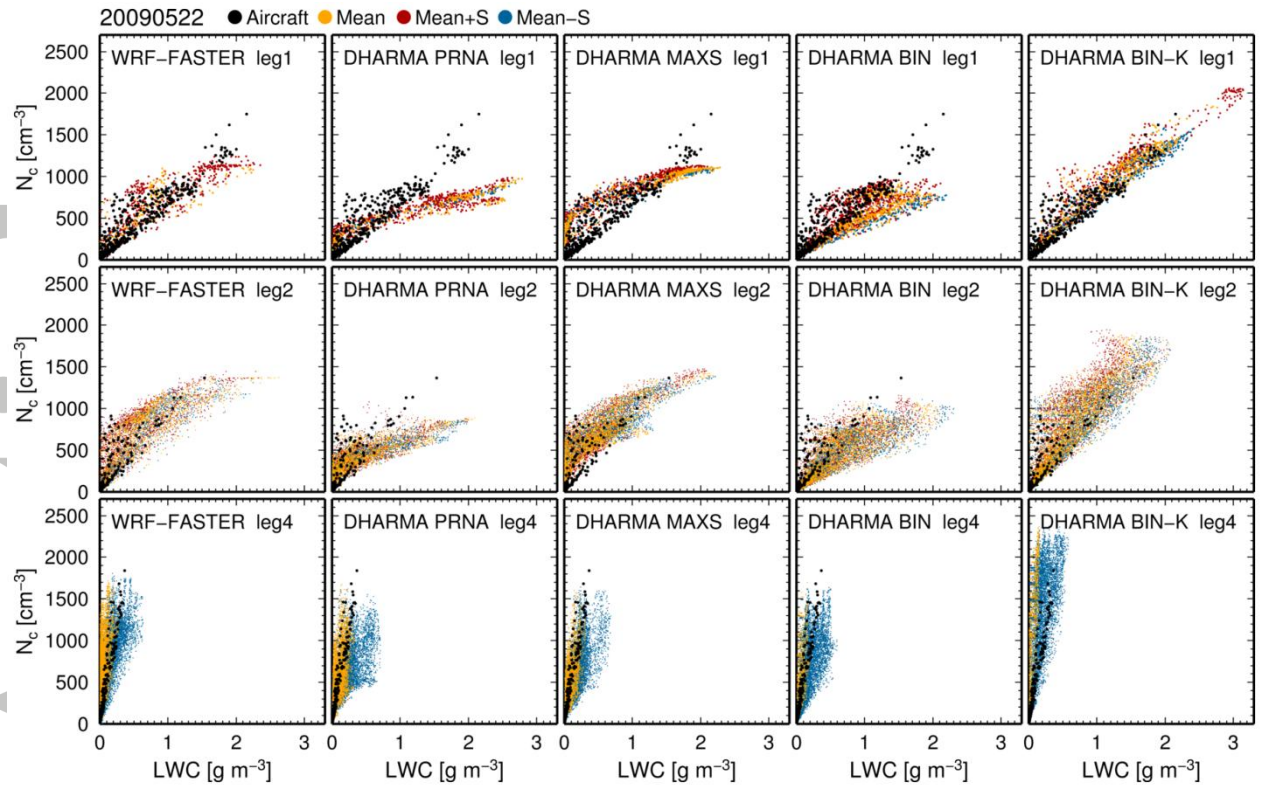


Figure 9: Scatter plots of liquid water content (LWC) and cloud droplet number concentration ( $N_c$ ) from simulations by WRF-FASTER, and DHARMA with the PRNA (baseline), MAXS, BIN, and BIN-K configurations (columns) for three horizontal flight legs (rows). Simulated clouds are sampled at the times and cloud-base-relative heights corresponding to the aircraft legs. Black dots indicate the aircraft measurements. Colors represent three altitude samplings of the simulations to consider the observational uncertainty in determining the relative height from cloud base. The cloud-base height in the observations is assumed to follow the time-dependent LCL height computed from the surface stations nearest to the flight track using their standard deviation as the uncertainty: mean (yellow), mean plus standard deviation (red), and mean minus standard deviation (blue). Legs 1 and 2 are sampled near cloud top and Leg 4 near cloud base. See text for further details.

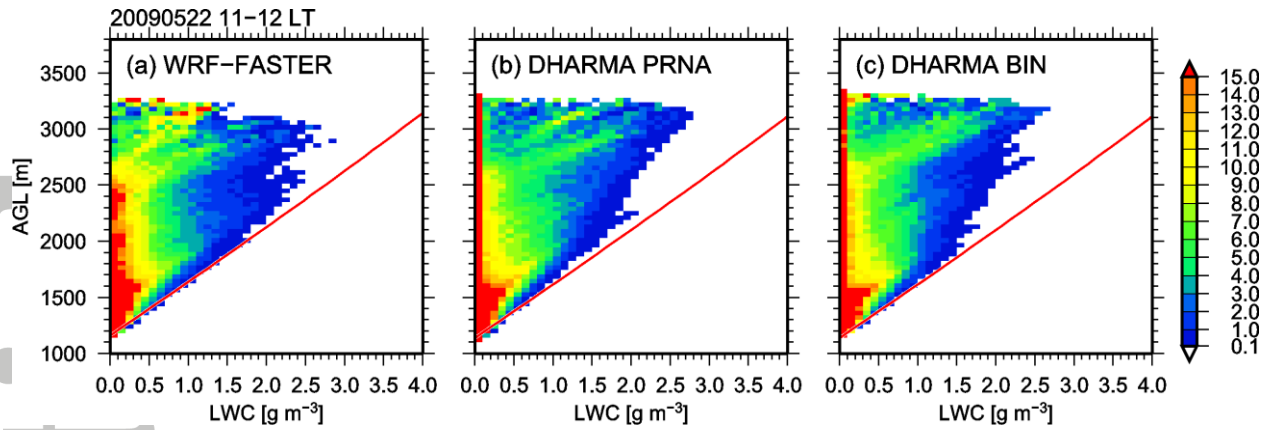


Figure 10: Height variations of the probability density functions of liquid water content (in percent using  $0.1 \text{ g m}^{-3}$  sized bins) from the simulations by WRF-FASTER (a), DHARMA PRNA (b), and DHARMA BIN (c). Red lines indicate the height variations of LWC for adiabatically-lifted parcels.



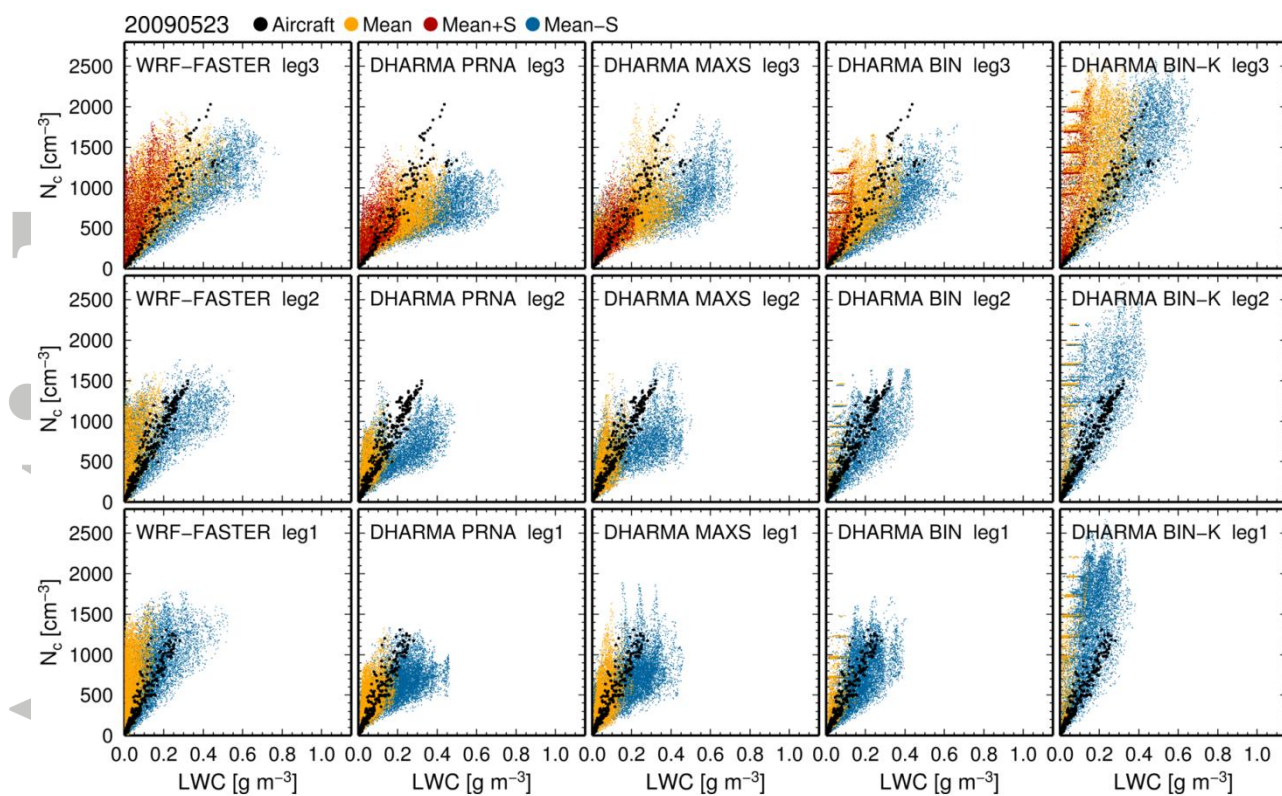


Figure 11: As in Fig. 9 but for the day 2 flight. Note that the x-scale is reduced from Fig. 9.

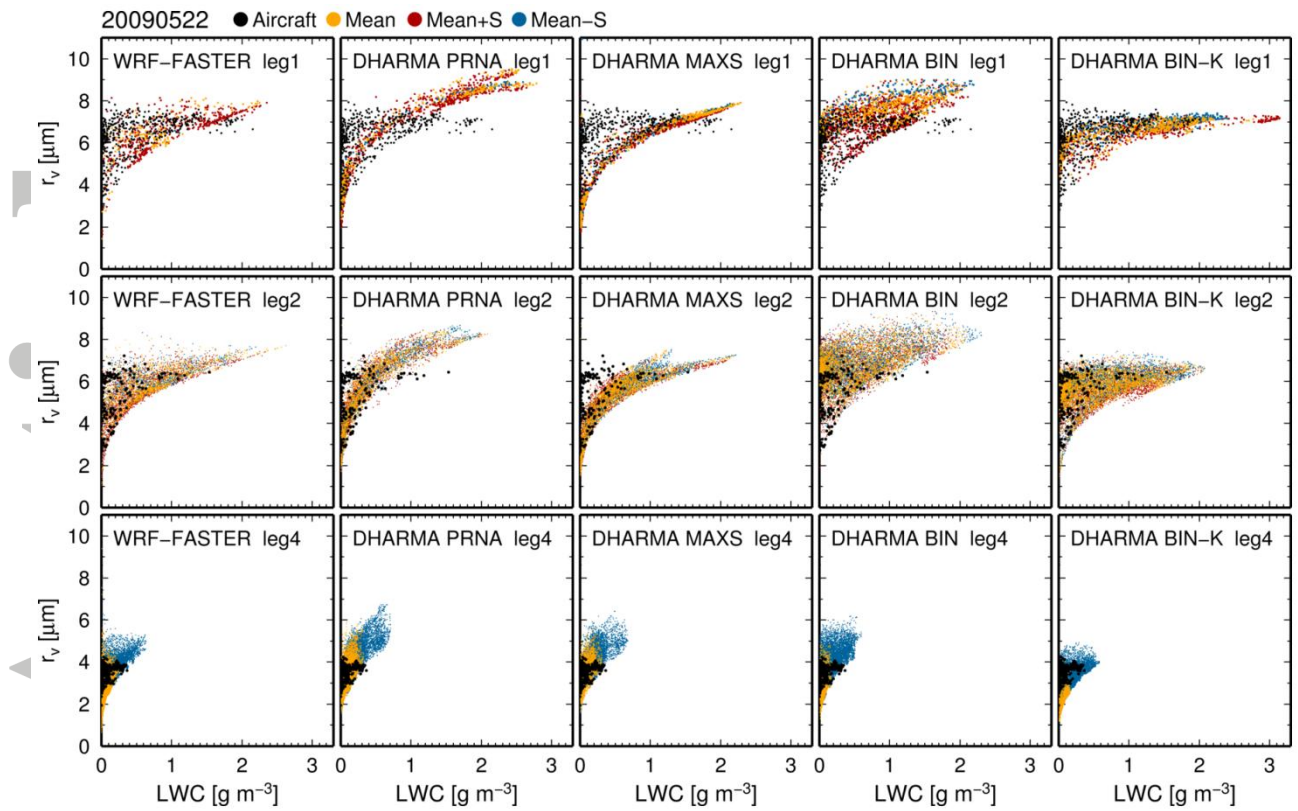


Figure 12: As in Fig. 9 but for liquid water content and volume mean radius for the day 1 flight.

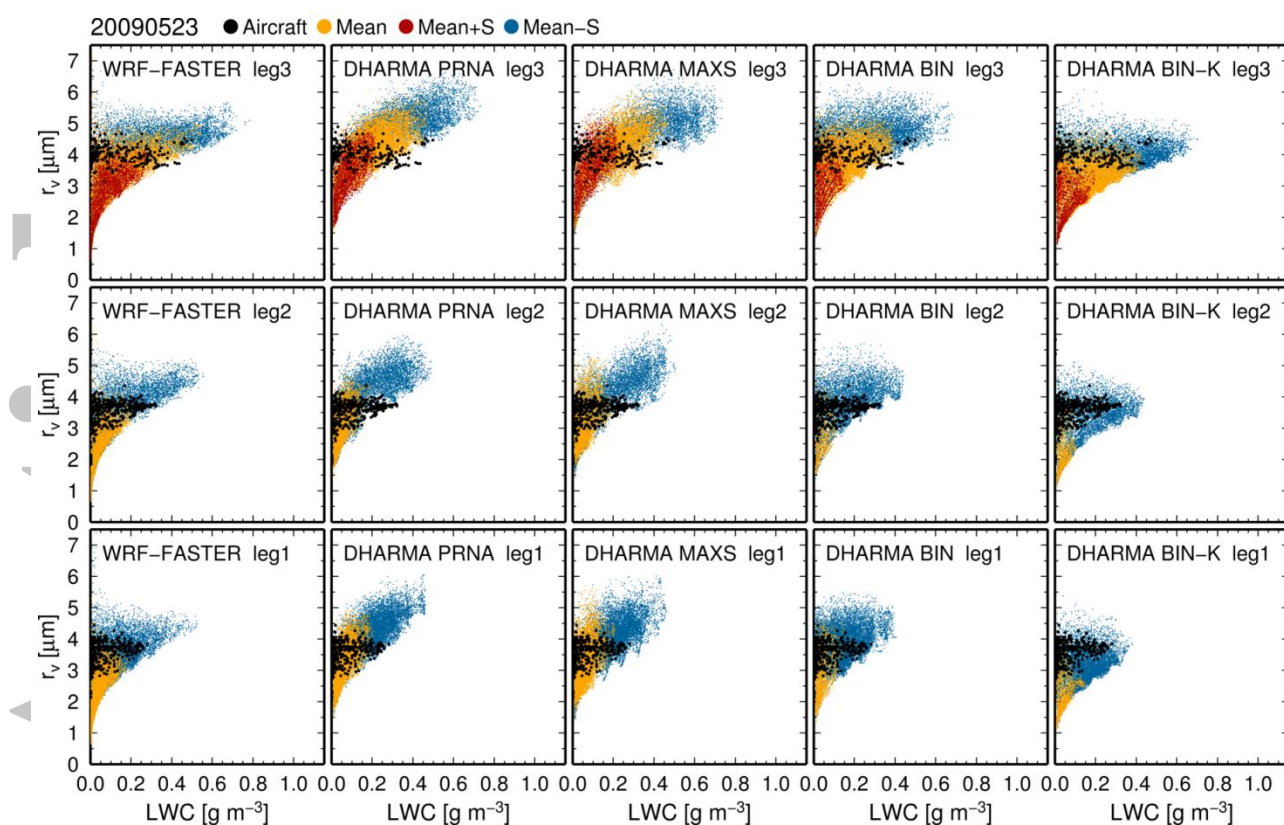


Figure 13: As in Fig. 12 but for liquid water content and volume mean radius for the day 2 flight.

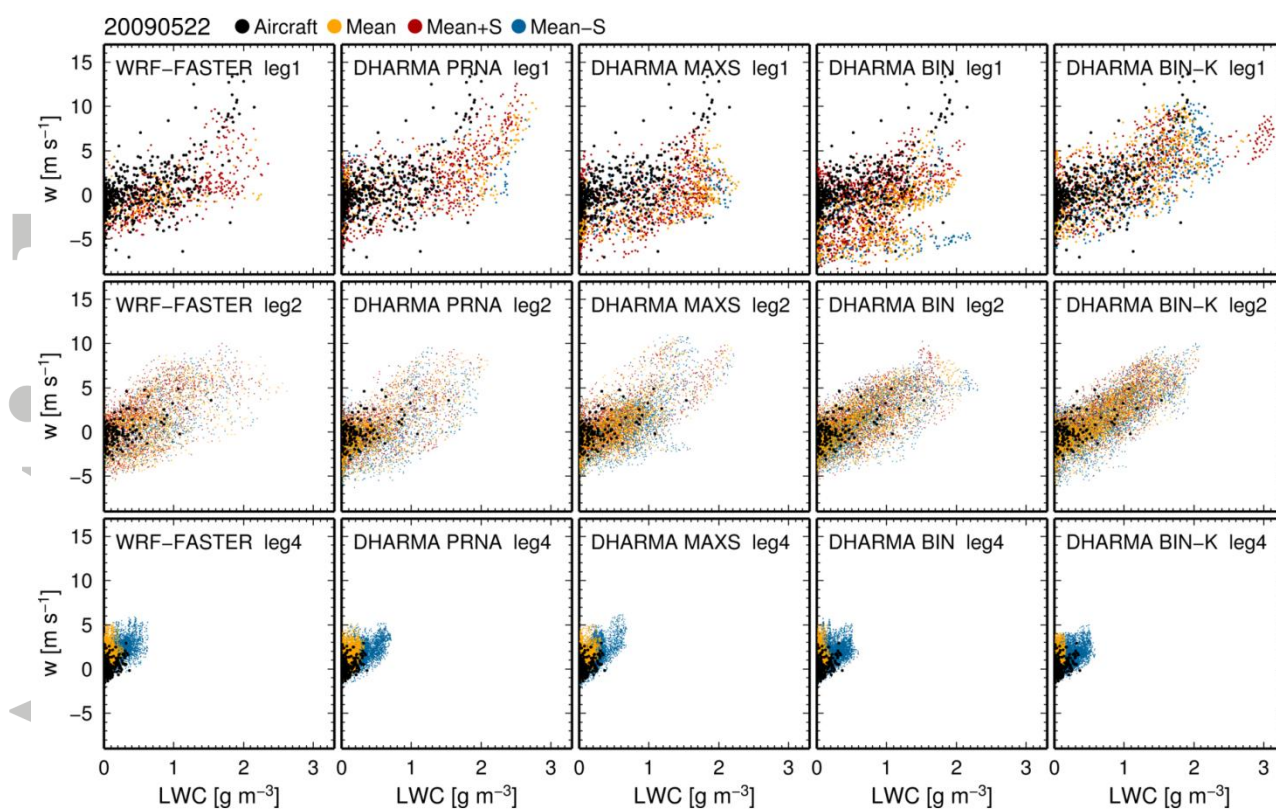


Figure 14: As in Fig. 9 but for liquid water content and vertical motion for the day 1 flight.



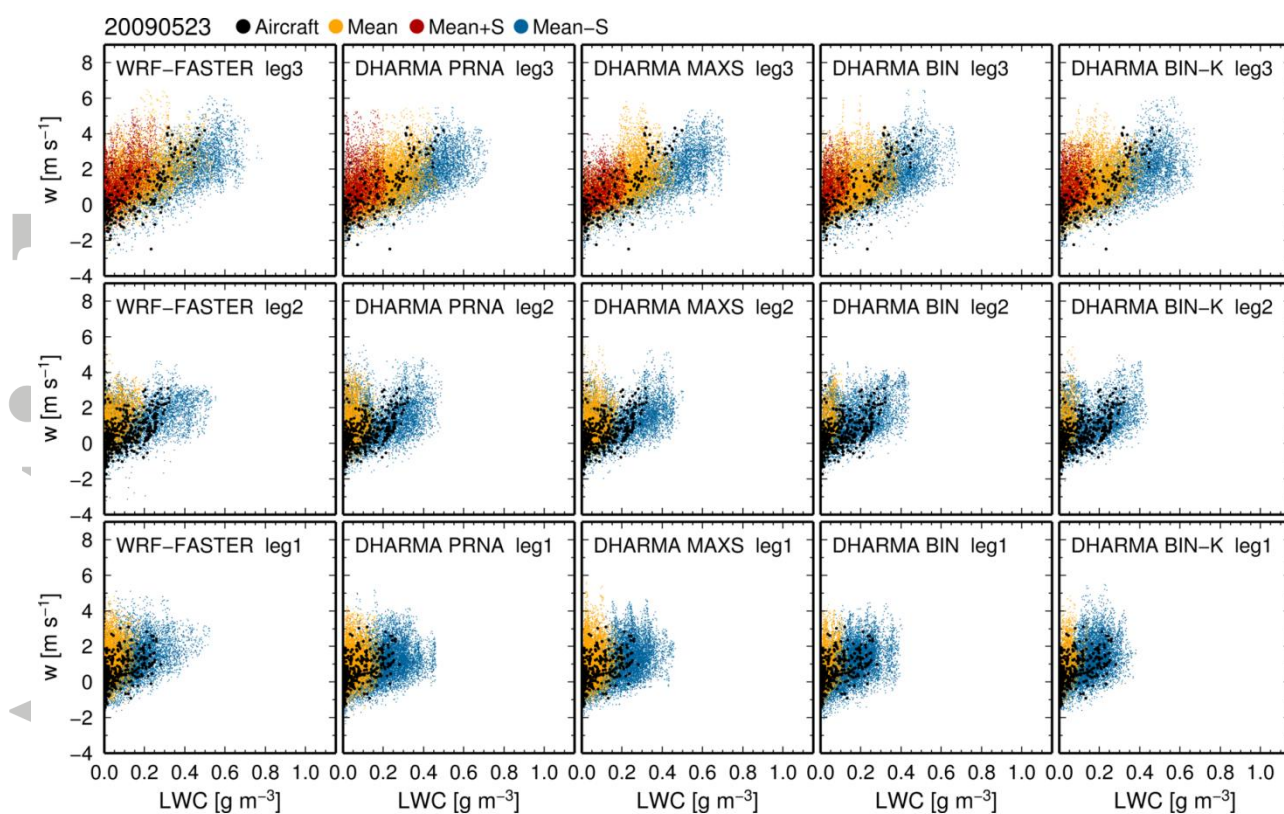


Figure 15: As in Fig. 14 but for liquid water content and vertical motion for the day 2 flight.

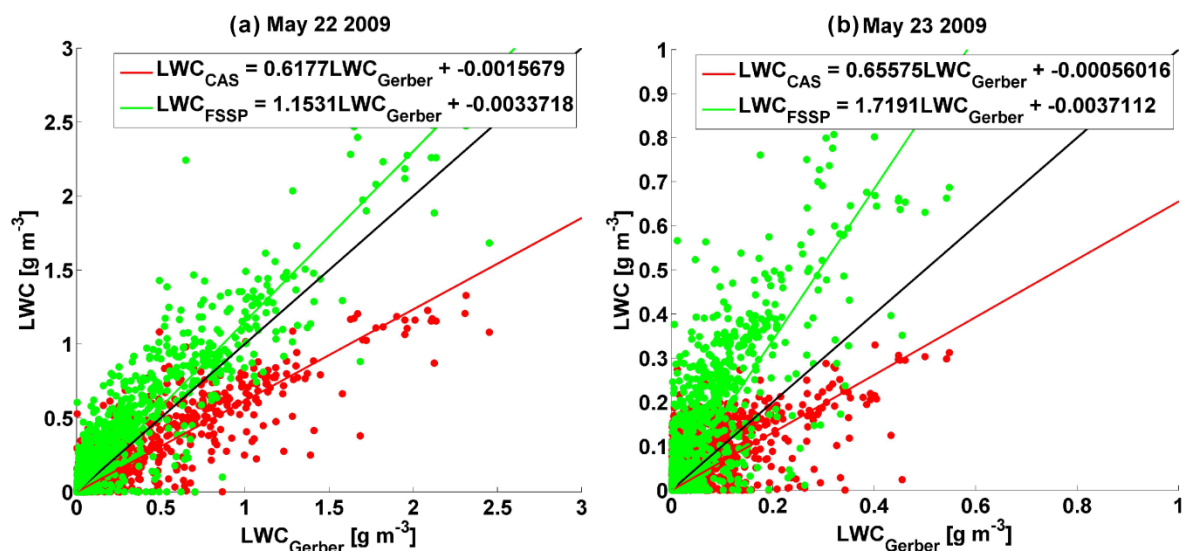


Figure B1: Scatter plots of liquid water content observed by Gerber Probe, CAS, and FSSP for the day 1 (a) and day 2 (b) flights. Red (green) indicates a comparison between Gerber Probe and CAS (FSSP).